目录

第 1 章

信号的时间-频率分析

信号实际上是传递信息的某种具体物理过程。最常用的信号分析方法是寻找一种简单有效的变换,使信号所包含的重要特征在变换域能更直接地显示出来。在小波变换兴起之前,傅立叶变换是信号分析中最重要的数学方法。傅立叶变换实际上是将信号展开为不同频率正弦信号的线性叠加。从信号的傅立叶变换,能看出信号各种不同频率成分的强弱,信号能量在频率域的分布。傅立叶变换的核函数是正弦函数,它在时间域上是无限的,非局部化的。而小波变换是将时间信号展开为小波函数族的线性叠加。小波变换的核函数是小波函数,它在时间域和频率域都是局部化的。这正是小波变换和传统傅立叶变换的本质区别。也正是因为小波变换在时间—频率域内都是局部化的,所以小波变换可对信号同时在时-频域内进行联合分析。

1.1 傅立叶变换

信号是一个或多个独立变量如时间、位置、距离、温度和压强等的函数。例如,语音和音乐信号表示为空间某一点的空气压强随时间变化的函数,一幅黑白图像表示为光的强度随两个空间坐标变化的函数。一维信号$f(t)$的傅立叶变换定义为

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{-i\omega t}\mathrm{d}t \tag{1.1.1}$$

傅立叶变换度量了信号在所有不同频率的振荡信息。傅立叶变换可以写成如下形式

$$F(\omega) = |F(\omega)|\mathrm{e}^{-i\theta(\omega)}$$

其中,$|F(\omega)|$称为信号$f(t)$的幅频特性,$\theta(\omega)$称为信号$f(t)$的相频特性。傅立叶变换的反变换为

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)\mathrm{e}^{i\omega t}\mathrm{d}\omega \tag{1.1.2}$$

意味着信号可展开为不同频率正弦信号的线性叠加。

由傅立叶变换的定义,很容易证明下述 Parseval 定理:

$$\int_{-\infty}^{+\infty} f(t)g^*(t)\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)G^*(\omega)\mathrm{d}\omega \tag{1.1.3}$$

令 $g(t)=f(t)$，则有

$$\int_{-\infty}^{+\infty}|f(t)|^2\mathrm{d}t=\frac{1}{2\pi}\int_{-\infty}^{+\infty}|F(\omega)|^2\mathrm{d}\omega \tag{1.1.4}$$

有时称上式为 Plancherel 定理。上式左边可解释为信号的能量，所以称 $|F(\omega)|^2$ 为信号的功率谱，它描述了信号能量在频率域内的分布。

根据傅立叶变换的定义，很容易证明傅立叶变换具有如下重要性质，它们有助于理解小波变换的物理意义，而且在若干定理的证明中也经常引用。

性　质	函　数	傅立叶变换
	$f(t)$	$F(\omega)$
逆	$F(t)$	$2\pi f(-\omega)$
卷积	$f_1(t)*f_2(t)$	$F_1(\omega)F_2(\omega)$
乘积	$f_1(t)f_2(t)$	$\frac{1}{2\pi}F_1(\omega)F_2(\omega)$
平移	$f(t-\tau)$	$F(\omega)e^{-i\tau\omega}$
调制	$f(t)e^{i\omega_0 t}$	$F(\omega-\omega_0)$
尺度伸缩	$f(t/s)$	$\|s\|F(s\omega)$
时域求导	$f^{(n)}(t)$	$(i\omega)^nF(\omega)$
频域求导	$(-it)^nf(t)$	$F^{(n)}(\omega)$
复共轭	$f^*(t)$	$F^*(-\omega)$

如信号是实的，即 $f(t)\in R$，即 $f^*(t)=f(t)$，由复共轭性质可得 $F(\omega)=F^*(-\omega)$，从而有

$$|F(\omega)|=|F(-\omega)|,\theta(\omega)=-\theta(-\omega)$$

也就是说，实信号的幅频特性是频率的偶函数，相频特性是频率的奇函数。

由于傅立叶变换的核函数 $e^{i\omega t}$ 在时域内是无限的，为了计算 $F(\omega)$，必须在信号的整个持续时间内积分。即为了获得信号中某一特定频率分量的信息，必须知道信号在整个时间过程中的变化情况。也就是说，傅立叶变换在时域内是非局部的。从上述分析还可以看到，时间函数 $f(t)$ 描述了信号的时域特征，其傅立叶变换 $F(\omega)$ 描述了信号的频域特征。也就是说，傅立叶变换要么在时域，要么在频域描述信号的特征，而不能对信号同时在时-频域内进行联合分析。

但在许多实际问题中，我们关心的却是信号在局部范围内的特征：例如音乐和语音信号中人们关心的是什么时刻演奏什么音符，发出什么样的音节；对地震波的记录人们关心的是什么位置出现什么样的反射波；图像处理中的边缘检测关心的是信号突变部分的位置，即纹理结构。尤其对非平稳信号的处理中，信号在任一时刻附近的频率特性都很重要。如柴油机缸盖表面的振动信号就是由撞击或冲击产生的，是一瞬变信号，仅从时域或频域上来分析都是不够的。R. Balian 认为："在通讯理论中，人们对于在完全给定的时间内，把一个振动信号表示成由其中每一个都拥有足够确定的位置与有一个频率的小波的叠加这件事感兴趣。事实上，有用的信息常常同时被发射信号的频率与信号的时间结构（如音乐）所传递。当把一个信号表

达成时间的函数时，其中的频谱表现并不好；相反地，信号的傅立分析却显示不了信号每一分量发射信号的瞬时与持续时间。一个适当的表示应结合这两者互补描述的优点，并用一个离散的刻画来表示，以适应通讯理论”。

针对瞬变的、非平稳信号的分析，曾经出现过许多改进的办法，其中比较有成效的有Wigner-Ville分布和窗口傅立叶变换两种，前者是一种非线性二次型变换。而D. Gabor于1946年提出的窗口傅立叶变换在非平稳信号的分析中起过很好的作用，而且与小波变换有很多相似之处。

1.2　窗口傅立叶变换

传统的傅立叶变换是对信号在整个时间过程中变化情况的分析，它在时域是全局化的，非局部的。其原因在于傅立叶变换的核函数正弦函数在时域是无限的。显然，将正弦函数乘以一个时域内衰减很快的函数之后，作为核函数是一种合理的改进方法，这种衰减很快的函数称为窗函数，以$g(t)$表示。一般来说，$g(t)$是一个实偶函数，且其能量主要集中在原点附近。例如均值为零的高斯函数，就是一个常用的窗函数。为了能实现对信号在整个时间过程中任意局部范围内变化情况的分析，应将窗函数沿时间轴移动，得$g(t-\tau)$，其能量主要集中在τ附近。对任意一个能量有限信号$f(t)\in L^2(R)$，其窗口傅立叶变换定义为

$$Gf(\omega,\tau)=\int_{-\infty}^{+\infty}f(t)g(t-\tau)e^{-i\omega t}dt \tag{1.2.1}$$

由于窗函数的局部性，所以上式中的积分实质上只在τ附近的一段时间内进行。也就是说，窗口傅立叶变换给出了信号在$t=\tau$附近的一段时间内的频率信息，窗口傅立叶变换在时域内是局部化的。

由上式可以明显地看到，窗口傅立叶变换确实是传统傅立叶变换的一种改进，即核函数由$e^{i\omega t}$改进为

$$g_{\omega,\tau}(t)=g(t-\tau)e^{i\omega t} \tag{1.2.2}$$

另一方面，也可以把窗口傅立叶变换理解为信号$f(t)$乘以窗函数$g(t-\tau)$后所得到的时间函数$f(t)g(t-\tau)$的傅立叶变换，所以窗口傅立叶变换又称为短时(Short-time)傅立叶变换。

根据希尔伯特空间中内积的定义，窗口傅立叶变换也可以写成内积形式

$$Gf(\omega,\tau)=\langle f(t),g_{\omega,\tau}(t)\rangle \tag{1.2.3}$$

应该注意，窗口傅立叶变换是频率ω和时间平移量τ的二元函数，所以信号$f(t)$的窗口傅立叶变换$Gf(\omega,\tau)$是对信号$f(t)$在时间-频率域内的联合分析。而信号$f(t)$的傅立叶变换$F(\omega)$仅是频率的函数，所以传统的傅立叶变换仅是对信号在频率域内的分析。

为了更清楚地了解窗口傅立叶变换在时间-频率域的性质，有必要先分析其核函数的性质。由于窗函数$g(t)$应选择为一平滑函数，所以其傅立叶变换$\hat{g}(\omega)$为一低通滤波器。由于在时间-频率平面上任意一点(ω_0,τ_0)处，核函数为

$$g_{\omega_0,\tau_0}(t)=g(t-\tau_0)e^{i\omega_0 t} \tag{1.2.4}$$

其傅立叶变换是

$$\hat{g}_{\omega_0,\tau_0}(\omega)=\hat{g}(\omega-\omega_0)e^{-i\tau_0(\omega-\omega_0)} \tag{1.2.5}$$

它显然是中心频率为 ω_0 的带通滤波器。

由(1.2.1)式可以得到在时间-频率平面内点(ω_0,τ_0)处的窗口傅立叶变换

$$Gf(\omega_0,\tau_0)=\int_{-\infty}^{+\infty}f(t)g^*_{\omega_0,\tau_0}(t)\mathrm{d}t=\int_{-\infty}^{+\infty}f(t)g(t-\tau_0)\mathrm{e}^{-i\omega_0 t}\mathrm{d}t \tag{1.2.6}$$

其中星号表示取共轭。应用 Parseval 定理,由上式可得:

$$\begin{aligned}Gf(\omega_0,\tau_0)&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(\omega)\hat{g}^*_{\omega_0,\tau_0}(\omega)\mathrm{d}\omega\\&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(\omega)\hat{g}(\omega-\omega_0)\mathrm{e}^{i(\omega-\omega_0)\tau_0}\mathrm{d}\omega\end{aligned} \tag{1.2.7}$$

由于$\hat{g}(\omega-\omega_0)$是中心频率为 ω_0 的带通滤波器,所以 $Gf(\omega-\tau_0)$可以解释为提取了信号在 $t=\tau_0$ 附近一段时间内在 $\omega=\omega_0$ 附近一段频率内的信息,也就是说,窗口傅立叶变换不仅在时域是局部化的,而且在频域也是局部化的。或者说,窗口傅立叶变换是对信号在时间-频率域内进行联合分析。为了更形象地说明这种联合分析,引入时-频窗的概念,如图 1.1 所示。时-频窗的中心位于(τ_0,ω_0)。将 $g(t)$规范化为$\|g(t)\|=1$,则$\|g_{\omega_0,\tau_0}(t)\|=1$。定义:

$$\begin{aligned}\sigma_t^2&=\int_{-\infty}^{+\infty}(t-\tau_0)^2|g_{\omega_0,\tau_0}(t)|^2\mathrm{d}t=\int_{-\infty}^{+\infty}t^2|g(t)|^2\mathrm{d}t\\\sigma_\omega^2&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}(\omega-\omega_0)^2|\hat{g}_{\omega_0,\tau_0}(\omega)|^2\mathrm{d}\omega=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\omega^2|\hat{g}(\omega)|^2\mathrm{d}\omega\end{aligned} \tag{1.2.8}$$

σ_t 和 σ_ω 分别称为 $g(t)$ 和 $\hat{g}(\omega)$的标准差,它们描述了图 1.1 所示时-频窗的宽度。对窗口傅立叶变换,其时-频窗的中心可以分别由(1.2.6)式和(1.2.7)式中的变量 τ_0 和 ω_0 来调节。但时-频窗的宽度与窗的中心(τ_0,ω_0)无关,是固定不变的,如(1.2.8)式所示。

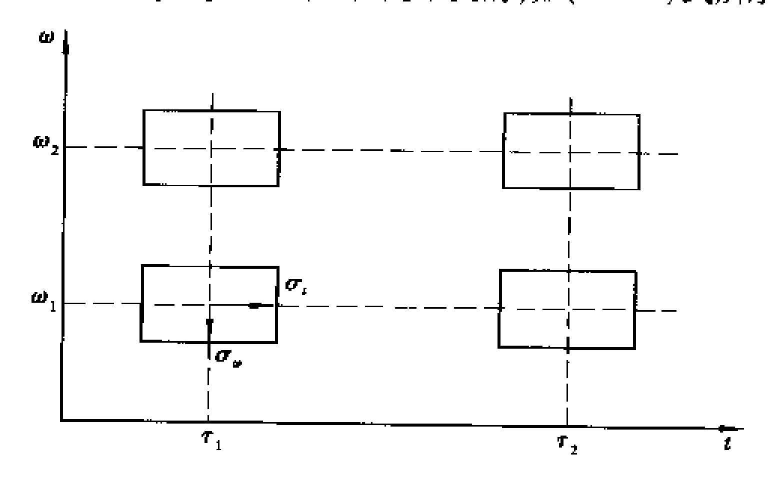


图 1.1 窗口傅立叶变换的时频窗

为了提高窗口傅立叶变换的精度,自然希望时-频窗的宽度在两个方向都同时变窄。但同时在时-频域内提高精度是有限制的,著名的测不准原理告诉我们

$$\sigma_t\sigma_\omega\geqslant\frac{1}{2} \tag{1.2.9}$$

当且仅当窗函数为高斯函数时,上式中的等号成立。和量子力学中的测不准原理一样,(1.2.9)式意味着人们不能无限地同时提高时-频域内的测试精度。这并不是由于测试仪器的限制,而是具有深刻的物理含义的。讨论某一时刻的频率为多少是没有意义的。

从数学的角度看,窗口傅立叶变换是将任一 $L^2(R)$空间中的函数$f(t)$映射为 $L^2(R^2)$空间

中的一个函数 $Gf(\omega,\tau)$。应用 Parseval 定理可以证明

$$\int_{-\infty}^{+\infty}|f(t)|^2\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|Gf(\omega,\tau)|^2\mathrm{d}\omega\mathrm{d}\tau \tag{1.2.10}$$

上式说明,经过窗口傅立叶变换之后,范数保持不变(带有一固定的比例系数),故窗口傅立叶变换具有惟一的反演公式

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}Gf(\omega,\tau)g(t-\tau)\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}\omega\mathrm{d}\tau \tag{1.2.11}$$

上式称为窗口傅立叶变换的反变换。应该注意,(1.2.10)隐含着如下意义:经过窗口傅立叶变换之后,并未丢失信息,只不过是将时间信号 $f(t)$ 所携带的信息在时间-频率域内用 $Gf(\omega,\tau)$ 表示出来,也正是因为窗口傅立叶变换未丢失信息,所以可以通过反变换将信号完整地恢复出来。

和傅立叶变换的功率谱类似,$|Gf(\omega,\tau)|^2$ 度量了信号在点(τ,ω)的邻域内的能量,该邻域的大小由时-频窗的宽度确定。

窗口傅立叶变换实现了信号在时间-频率域内局部化的联合分析,但是其时间-频率窗的宽度却是固定不变的。当然可以选择不同类型的窗函数,或者调整窗函数的参数(例如可以调整高斯函数的方差),来选择时间-频率窗的宽度。但是一旦窗函数的类型及参数确定之后,惟一可以调整的就是平移量 τ,而变量 τ 只能改变窗沿时间轴的位置,不能改变时-频窗的宽度。窗口傅立叶变换中的变量 ω 也仅仅是改变窗沿频率轴的位置,对时-频窗的宽度没有影响。这样窗口傅立叶变换就不能适应许多复杂的信号分析和处理。上述这些性质都是由窗口傅立叶变换的核函数决定的。

1.3　连续小波变换

下面将要详细讨论的连续小波变换也是信号在时间-频率域内局部化联合分析的一种方法。小波 $\psi(t)$ 是一个时间函数,它正负波动的同时速降至零,其傅立叶变换 $\hat{\psi}(\omega)$ 呈现为带通滤波器的频率特性,也就是说,小波在时域和频域都是局部化的。将 $\psi(t)$ 伸缩和平移而得到的一族函数 $\psi_{a,b}(t)$ 称为分析小波,它在时域和频域也是局部化的。以 $\psi_{a,b}(t)$ 为核函数的积分变换,正是本节要讨论的积分小波变换或称连续小波变换。与窗口傅立叶变换不同的是,小波变换的时间-频率窗不是固定不变的,而是可以自适应调整的。这是小波变换和窗口傅立叶变换的根本区别。也正是由于这个特点,决定了小波分析在实际应用中的独特地位。

1.3.1　连续小波变换

小波是一种持续时间很短的波,但并不是任意持续时间很短的波都是小波,小波必须满足以下容许条件

$$C_\psi = \int_0^{+\infty}\frac{|\hat{\psi}(\omega)|^2}{\omega}\mathrm{d}\omega < +\infty \tag{1.3.1}$$

满足上式的时间函数 $\psi(t)$ 称为母小波或基本小波,通常简称为小波。容许条件意味着 $\hat{\psi}(\omega)$ 在 $\omega=0$ 处的值必须为零,即 $\hat{\psi}(0)=0$,从傅立叶变换的定义可得

$$\int_{-\infty}^{+\infty}\psi(t)\mathrm{d}t = 0 \tag{1.3.2}$$

由上式可以清楚地看到，$\psi(t)$必须时正时负地波动，否则，$\psi(t)$的积分不会为零。在实际应用中，要求小波具有很好的时域局部化性质，即当$t\to\pm\infty$时，要求$\psi(t)$速降至零。也就是说，小波是持续时间很短的衰减振荡，它在时域内是局部的。正弦波也时正时负地波动，但是它是一种等幅的波动，在时间上是无限的，它显然不满足(1.3.2)式。“小波”不是指其波动的幅度很小，而是指其持续时间很短。在频率域，同样也要求当$\omega\to\pm\infty$时，$\hat{\psi}(\omega)$速降至零，这也是容许条件(1.3.1)式对$\hat{\psi}(\omega)$的约束。再注意到$\hat{\psi}(0)=0$，可见$\hat{\psi}(\omega)$为一带通滤波器的频率特性。这就是说，小波在频域内也是局部的。小波在时域和频域内都是局部的，这是小波最重要的特性。小波在时-频域内的局部性实际上是其能量在时-频域内的集中性。

为了定量描述小波在时-频域内的局部性，定义小波函数的中心t_0和标准差σ_{t_0}

$$t_0 = \int_R t|\psi(t)|^2\mathrm{d}t/\|\psi(t)\|^2$$

$$\sigma_{t_0}^2 = \int_R (t-t_0)^2|\psi(t)|^2\mathrm{d}t/\|\psi(t)\|^2 \tag{1.3.3}$$

其傅立叶变换$\hat{\psi}(\omega)$的中心ω_0和标准差σ_{ω_0}分别定义为

$$\omega_0 = \int_{R+}\omega|\hat{\psi}(\omega)|^2\mathrm{d}\omega/\|\hat{\psi}(\omega)\|^2$$

$$\sigma_{\omega_0}^2 = \int_{R+}(\omega-\omega_0)^2|\hat{\psi}(\omega)|^2\mathrm{d}\omega/\|\hat{\psi}(\omega)\|^2 \tag{1.3.4}$$

t_0和σ_{t_0}分别度量小波的中心和时间宽度，ω_0和σ_{ω_0}分别度量带通滤波器$\hat{\psi}(\omega)$的中心频率和带宽。

将母小波伸缩和平移之后得到的函数族

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \qquad a\in R^+, b\in R \tag{1.3.5}$$

称为分析小波。式中a为伸缩参数或尺度参数，取正实数。当$a>1$时，沿时间轴方向拉伸。当$a<1$时，沿时间轴方向压缩。因子$1/\sqrt{a}$是为保持伸缩之后能量不变，或者说范数不变，即$\|\psi_{a,b}(t)\|=\|\psi(t)\|$。$b$为平移参数，可以取任意实数。实际上，窗口傅立叶变换中不同频率可理解为正弦函数疏密程度的不同，也可以理解为是沿时间轴的伸缩。只不过窗口傅立叶变换中的伸缩和平移是分别针对正弦函数和窗函数的，而小波变换中伸缩和平移都是针对小波函数的。图1.2中，中间波形为母小波$\psi(t)$，其对称中心位于0；左面为$a=2, b=-8$时的分析小波；右面为$a=1/2, b=5$时的分析小波。

由图1.2可以清楚地看到，分析小波在时域是局部化的。按照类似的定义，分析小波$\psi_{a,b}(t)$的中心t'和标准差σ_t'为

$$t' = at_0 + b, \sigma_t' = a\sigma_{t_0} \tag{1.3.6}$$

分析小波$\psi_{a,b}(t)$的傅立叶变换$\hat{\psi}_{ab}(\omega)=\sqrt{a}\hat{\psi}(a\omega)e^{-ib\omega}$，它仍然是一个带通滤波器，也就是说，分析小波在频域也是局部化的，其中心和标准差分别为

$$\omega' = \frac{\omega_0}{a}, \sigma_\omega' = \frac{\sigma_{\omega_0}}{a} \tag{1.3.7}$$

改变尺度参数意味着带通滤波器的中心频率ω'和带宽$2\sigma_\omega'$都在调整，但带通滤波器的Q值

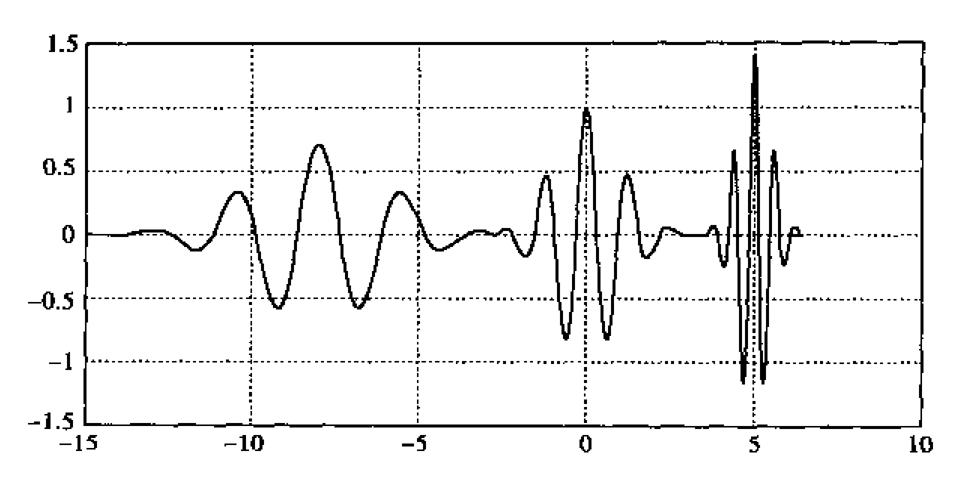


图 1.2　分析小波

$2\sigma'_{\omega}/\omega'$ 却是恒定不变的。

对任意能量有限信号 $f(t) \in L^2(\mathrm{R})$，其连续小波变换定义为

$$Wf(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)\psi^*\left(\frac{t-b}{a}\right)\mathrm{d}t = \langle f(t), \psi_{a,b}(t)\rangle \tag{1.3.8}$$

式中星号表示取共轭。可以看到，连续小波变换和传统的傅立叶变换及窗口傅立叶变换其数学描述方法是类似的，都是取信号和核函数的内积。这些变换也都可以解释为信号和核函数相关程度的度量。而各种变换的区别在于选取的核函数不同。从数学形式上看，傅立叶变换将一维时间函数 $f(t)$ 映射为一维频率函数 $F(\omega)$，所以傅立叶变换是对时间信号的频率分析。窗口傅立叶变换将一维时间函数 $f(t)$ 映射为二维函数 $Gf(\omega,\tau)$，是对时间信号的时-频联合分析。类似的，小波变换将一维时间函数 $f(t)$ 映射为二维函数 $Wf(a,b)$，也是对时间信号的时-频联合分析。由于连续小波变换是按积分形式定义的，所以又称为积分小波变换。由于 $\psi_{a,b}(t)$ 是局部化的，所以连续小波变换也是对信号在时间域内的局部化分析，它度量了信号在某个邻域的变化，该邻域以 t' 为中心，宽度与 a 成正比。实际上，当尺度趋向于零时，小波变换的衰减性刻画了 $f(t)$ 在 t' 的邻域内的正则性。这一性质在瞬变检测、分形分析中有重要的应用。

如定义

$$\bar{\psi}_a(t) = \frac{1}{\sqrt{a}}\psi^*\left(\frac{-t}{a}\right) \tag{1.3.9}$$

则连续小波变换可以写成卷积形式

$$Wf(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)\psi^*\left(\frac{t-b}{a}\right)\mathrm{d}t = f * \bar{\psi}_a(b) \tag{1.3.10}$$

也就是说，连续小波变换可以理解成信号通过冲击响应为 $\bar{\psi}_a(t)$ 的系统，该系统的频率特性为

$$\hat{\bar{\psi}}_a(\omega) = \sqrt{a}\hat{\psi}^*(a\omega)$$

它是一个带通滤波器，其中心频率和带宽分别为 ω_0/a 和 $2\sigma_{\omega_0}/a$。于是，小波变换也可理解为信号通过带通滤波器。

现在可以清楚地看到，连续小波变换确实是对信号在时-频域内的联合分析。所以仍然可以用时-频窗来形象地表示小波变换的局部化，如图 1.3 所示。尺度 a 决定了时域和频域观测窗的大小，即观测的范围或者说尺度。而变量 b 只影响时域内窗的位置。所以积分小波变换

也可以说是对信号的时间-尺度分析,这样说或许更准确一些。当然尺度和频率确实存在着一定的联系,但在许多实际应用中这种联系并不是那么明显。

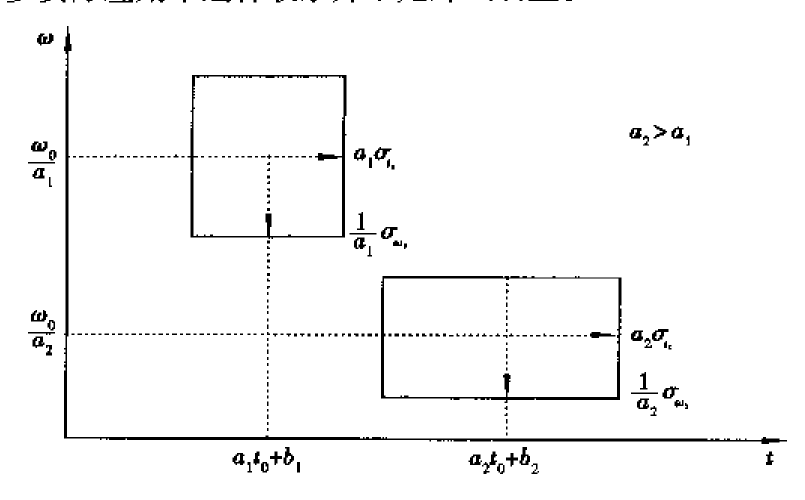


图 1.3　小波变换的时频窗

在小波变换中,尺度 a 越大,时窗越宽,频窗越窄且频窗中心往低频方向移动;a 越小,则时窗越窄,频窗越宽且频窗中心向高频方向移动。也就是说,为了获得信号的低频信息,应做较长时间的观察,这时有较好的频域分辨率。而做短时间的观察时,将获得信号的高频信息,这时有较好的时域分辨率,这正是小波分析的优良性质。人们将小波变换的这种性质形象地比喻为"数学显微镜"。大范围地观察时,看到的是被测物的大致全貌,这时提取的是低频信息。而为了看到细节(高频信息),则必须将镜头聚焦于感兴趣的部分。所以也有人将它称为小波变换的"变焦"特性。由于带通滤波器的 Q 值是恒定的,故小波变换也相当于一组等 Q 滤波器。中心频率较高时频率分辨率较低,而中心频率较低时则有较高的频率分辨率。

当时间-尺度连续变化时,分析小波的时间-频率窗覆盖整个时间-频率平面。可以预见,f 可以由它的小波变换 $Wf(a,b)$ 来恢复。在实际应用中,大多数小波都是实函数。如 $\psi\in L^2(R)$ 且满足容许条件(1.3.1)式,则可以证明:对任意 $f\in L^2(R)$,均有

$$f(t) = \frac{1}{c_\psi}\int_0^{+\infty}\frac{\mathrm{d}a}{a^2}\int_{-\infty}^{+\infty}Wf(a,b)\psi_{a,b}(t)\,\mathrm{d}b \tag{1.3.11}$$

及

$$\int_{-\infty}^{+\infty}|f(t)|^2\mathrm{d}t = \frac{1}{c_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty}|Wf(a,b)|^2\frac{\mathrm{d}a\mathrm{d}b}{a^2} \tag{1.3.12}$$

上式左边是信号的能量,故其小波变换的模的平方表示了信号能量在时间-尺度平面内的分布,因而 $|Wf(a,b)|^2$ 称为信号的尺度谱。

在 Matlab 的小波工具箱中,函数 cwt 可以实现连续小波变换。图 1.4(a)、(b)分别表示出受噪音污染的方块信号及其连续小波变换。由于一维信号的连续小波变换是尺度参数 a 和时间平移参数 b 的二维函数,所以图 1.4(b)中横轴为 b,纵轴为 a,小波变换的模用灰度表示。在信号的平坦部分,其小波变换的模很小,故灰度等级很低。这一点也很容易从小波变换的定义(1.3.8)式看出,当信号在小波的持续时间内不变或变化很缓慢时,$f(t)$ 可从积分号内移出,由于分析小波的积分为零,故此时其小波变换的值为零或很小。在信号的突变部分,其小波变换的模很大,故灰度等级很高,对应图中的亮条纹。由这个简单的例子不难看到小波分析的时-频局部性。同时应注意,不同的(a,b)值一般对应不同的小波变换值 $Wf(a,b)$,但这些不同

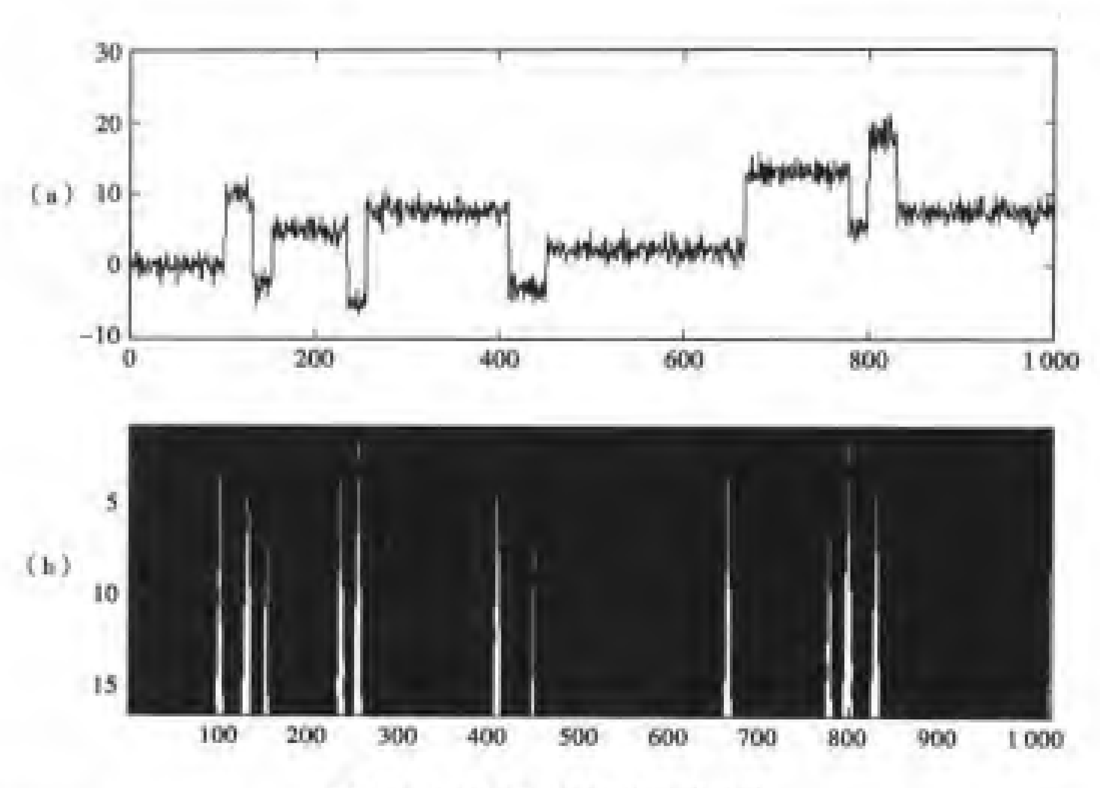


图1.4　(a)信号,(b)连续小波变换

的小波变换值 $Wf(a,b)$之间有一定的相关性,这就增加了分析和解释小波变换的困难。因此,小波变换的冗余度应尽可能减小,它是小波分析中的主要问题之一。

还应当注意,傅立叶变换用到的核函数为正弦函数,具有惟一性;小波分析用到的核函数为小波函数则有很多种,具有不惟一性,同一工程问题用不同的小波函数进行分析有时结果相差甚远。小波函数的选用是小波分析实际应用中的一个难点,也是小波分析研究的一个热点,目前往往是根据经验或通过不断的试验,对结果进行对比分析来选择合适的小波。在1.5节中,将介绍几种经常使用的小波。

1.3.2　连续小波变换的性质[9,11]

连续小波变换具有如下性质

性质1(线性)　设 $f(t)=\alpha g(t)+\beta h(t)$,则

$$Wf(a,b)=\alpha Wg(a,b)+\beta Wh(a,b)$$

性质2(平移不变性)　若 $f(t)\leftrightarrow Wf(a,b)$,则 $f(t-\tau)\leftrightarrow Wf(a,b-\tau)$。平移不变性是一个很好的性质,在实际应用中,尽管离散小波变换要用得广泛一些,但在需要有平移不变性的情况下,离散小波变换是不能直接使用的。

性质3(伸缩共变性)　若 $f(t)\leftrightarrow Wf(a,b)$,则 $f(ct)\leftrightarrow \frac{1}{\sqrt{c}}Wf(ca,cb)$,其中 $c>0$。

性质4(内积定理)　若 $f(t)\leftrightarrow Wf(a,b)$, $g(t)\leftrightarrow Wg(a,b)$则有

$$\langle f(t),g(t)\rangle=\frac{1}{c_\psi}\langle Wf(a,b),Wg(a,b)\rangle \tag{1.3.13}$$

式中 c_ψ 由容许条件(1.3.1)式定义。内积定理也叫 Moyal 定理。右边的内积是对 a 和 b 的双重积分,而且由于小波变换定义式中 a 以倒数形式出现,所以微分为 da/a^2。这样便可写出内

积定理更具体的形式

$$\int_{-\infty}^{+\infty} f(t) g^{*}(t) \mathrm{d}t = \frac{1}{c_{\psi}} \int_{0}^{+\infty} \frac{\mathrm{d}a}{a^{2}} \int_{-\infty}^{+\infty} Wf(a,b) Wg^{*}(a,b) \mathrm{d}b \tag{1.3.14}$$

如令 $g(t) = f(t)$，则可由上式得到能量不变性表达式(1.3.12)。

性质5(冗余性) 连续小波变换中存在信息表述的冗余度。尽管冗余的存在可以提高信号重建时计算的稳定性，但增加了分析和解释小波变换结果的困难。(1.3.11)式意味着 $f(t)$ 可分解成 $\psi_{a,b}(t)$ 的线性组合，其系数就是 $f(t)$ 的小波变换。由于 $\psi_{a,b}(t)$ 的参数 a,b 是连续变化的，所以 $\psi_{a,b}(t)$ 之间不是线性无关的，也就是说，它们之中有"富余"的，这就导致 $Wf(a,b)$ 之间有相关性。具体来说，

$$\begin{aligned} Wf(a_0,b_0) &= \int_{-\infty}^{+\infty} f(t) \psi_{a_0,b_0}^{*}(t) \mathrm{d}t \\ &= \int_{-\infty}^{+\infty} \frac{1}{c_{\psi}} \left(\int_{0}^{+\infty} \int_{-\infty}^{+\infty} Wf(a,b) \psi_{a,b} \frac{\mathrm{d}a\mathrm{d}b}{a^{2}} \right) \psi_{a_0,b_0}^{*}(t) \mathrm{d}t \\ &= \frac{1}{c_{\psi}} \int_{0}^{+\infty} \int_{-\infty}^{+\infty} Wf(a,b) \left(\int_{-\infty}^{+\infty} \psi_{a,b}(t) \psi_{a_0,b_0}^{*}(t) \mathrm{d}t \right) \frac{\mathrm{d}a\mathrm{d}b}{a^{2}} \\ &= \frac{1}{c_{\psi}} \int_{0}^{+\infty} \int_{-\infty}^{+\infty} Wf(a,b) K(a,a_0,b,b_0) \frac{\mathrm{d}a\mathrm{d}b}{a^{2}} \end{aligned} \tag{1.3.15}$$

其中

$$K(a,a_0,b,b_0) = \langle \psi_{a,b}(t), \psi_{a_0,b_0}(t) \rangle \tag{1.3.16}$$

$K(a,a_0,b,b_0)$ 称为再生核，它度量了两个分析小波 $\psi_{a,b}(t)$ 和 $\psi_{a_0,b_0}(t)$ 的相关性。(1.3.15)式表明：在 (a,b) 平面上一个特定的点 (a_0,b_0) 的小波变换 $Wf(a_0,b_0)$ 都要受到其他各处的小波变换 $Wf(a,b)$ 的"影响"，再生核度量了"影响"程度的大小。如 $\psi_{a,b}(t)$ 和 $\psi_{a_0,b_0}(t)$ 正交，再生核为零，$Wf(a,b)$ 对 $Wf(a_0,b_0)$ 没有影响。

由于连续小波变换的冗余性，并不是每个 $\Phi(a,b) \in L^2(R^2)$ 都是某个 $L^2(R)$ 函数 f 的小波变换，当且仅当 $\Phi(a,b)$ 满足如下再生核方程

$$\Phi(a_0,b_0) = \frac{1}{c_{\psi}} \int_{0}^{+\infty} \int_{-\infty}^{+\infty} K(a,a_0,b,b_0) \Phi(a,b) \frac{\mathrm{d}a\mathrm{d}b}{a^{2}} \tag{1.3.17}$$

时，$\Phi(a,b)$ 才是某个 $L^2(R)$ 函数 f 的小波变换。

1.4 离散小波变换

在连续小波变换中，a 和 b 是连续变化的，时频窗在时频平面是连续移动的，所以连续小波变换存在信息冗余。不难理解，在一定条件下，a 和 b 可以离散化而不丢失信息。其中最有意义的是 a 按2的整数次幂变化，即 $a = 2^j, j \in Z$，b 应与 a 成正比，故取 $b = 2^j n, n \in Z$，即按步长 2^j 作整数平移。尺度参数和平移参数离散化之后的分析小波写成

$$\psi_{j,n}(t) = 2^{-\frac{j}{2}} \psi(2^{-j} t - n); \quad j,n \in Z \tag{1.4.1}$$

其中，j 为伸缩参数或尺度参数，n 为沿时间轴的平移参数，$j < 0$ 表示将 $\psi(t)$ 沿时间轴压缩，尺度更精细，平移步长小，$j > 0$ 表示将 $\psi(t)$ 沿时间轴拉伸，尺度更粗糙，平移步长大。

将(1.4.1)式代入连续小波变换表达式(1.3.8)得

$$c_{j,n} = \langle f(t), \psi_{j,n}(t) \rangle = \int_{-\infty}^{+\infty} f(t)\psi_{j,n}^{*}(t)\,\mathrm{d}t$$

$c_{j,n}$实际上是连续小波变换 $Wf(a,b)$在时间-尺度平面的离散网格上的抽样,那么经过抽样以后是否丢失了信息呢? 换句话说,是否可以从 $c_{j,n}$重构$f(t)$呢? 这个问题将在框架理论中仔细讨论。现在只讨论一种最简单但却最重要的情况,即母小波 $\psi(t)$经二进伸缩和平移得到的函数族$\{\psi_{j,n}(t)\}_{j,n\in Z}$构成 $L^2(R)$的正交规范基。

如小波函数族构成 $L^2(R)$的正交规范基,即

$$\langle \psi_{j,n}(t), \psi_{l,k}(t) \rangle = \delta_{j,l}\delta_{n,k} \tag{1.4.2}$$

则任意$f(t) \in L^2(R)$可展开为正交规范基的线性组合

$$f(t) = \sum_{j,n} c_{j,n}\psi_{j,n}(t) \tag{1.4.3}$$

由(1.4.2)式很容易求得展开式的系数

$$c_{j,n} = \langle f(t), \psi_{j,n}(t) \rangle = \int_{-\infty}^{+\infty} f(t)\psi_{j,n}^{*}(t)\,\mathrm{d}t \tag{1.4.4}$$

上式定义了函数$f(t)$的离散小波变换,而式(1.4.3)则是离散小波变换的反变换,$c_{j,n}$称为$f(t)$的离散小波变换或离散小波系数。通过 2 的整数次幂伸缩和整数平移构成正交规范基的小波称为正交小波。必须注意,并不是所有的小波都是正交小波。

由(1.4.2)、(1.4.3)和(1.4.4)式不难证明

$$\langle f(t), f(t) \rangle = \int_{-\infty}^{+\infty} |f(t)|^2\mathrm{d}t = \sum_{j,n} |c_{j,n}|^2 \tag{1.4.5}$$

这就是离散小波变换的 Parseval 定理。如引入范数的定义

$$\|f(t)\| = \sqrt{\langle f(t), f(t) \rangle}, \|c_{j,n}\| = \sqrt{\langle c_{j,n}, c_{j,n} \rangle} = \sqrt{\sum_{j,n} |c_{j,n}|^2}$$

则(1.4.5)式也可写成

$$\|f(t)\| = \|c_{j,n}\| \tag{1.4.6}$$

这实际上就是离散小波变换的能量恒等式。也就是说,离散小波变换是保持范数不变的映射。如时间信号经过变换之后,在时间域的范数与变换域的范数相等(或带一常系数),则变换未丢失信息,存在反变换。在正交基的条件下,变换也不存在信息冗余。放松正交性的要求意味着放松能量恒等条件,在第 2 章框架理论中将详细讨论这个问题。

需要强调的是,离散小波变换所分析的信号仍然是连续时间信号,只是尺度参数和时间平移参数离散化而已,这和离散傅立叶变换是不同的,希望不要因为名称上的类似而产生概念上的混淆。另外也不要因为正交展开式在数学形式上的类似,把离散小波变换和傅立叶级数在物理概念上混淆起来。傅立叶级数是对周期性时间信号的频谱分析,而离散小波变换并不要求信号是周期性的。

离散小波变换同样也是对信号的时-频局部化分析,离散小波系数 $c_{j,n}$中j表示尺度或者说频率,而 n 表示时间。由于小波在时-频域都是局部化的,所以对那些只有少量快速变化点或区域的信号而言,只有少量的系数较大,而其他系数都很小甚至为零,尤其对暂态的,非平稳的信号更是如此。这也是小波分析成功应用于数据压缩、去噪和信号检测的根本原因。

对于离散小波变换的计算,存在快速算法,即第 3 章将要介绍的 Mallat 算法。在 Matlab 的

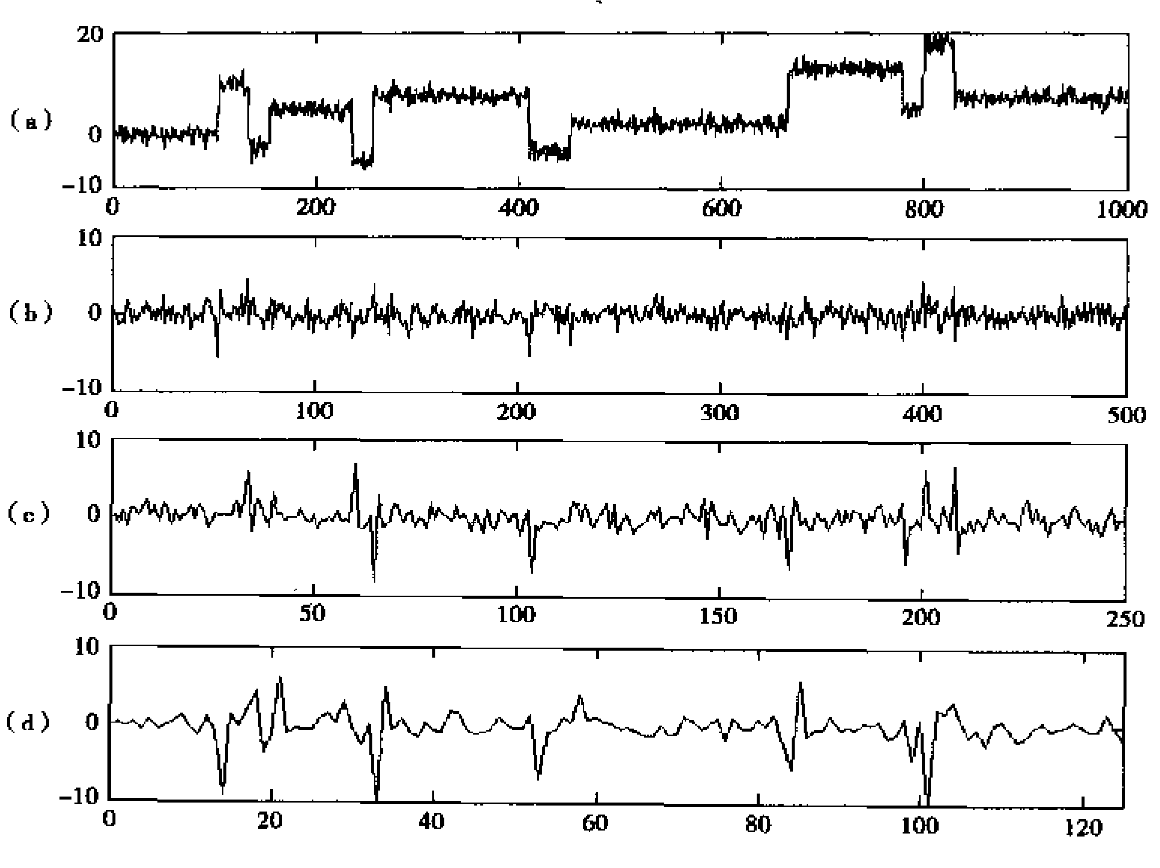


图 1.5　离散小波变换

小波工具箱中，函数 dwt 可以用来实现离散小波变换，但是它只能实现原始信号的单级分解。原始信号可能是实际模拟信号的离散抽样，也可能它本来就是离散的数字序列。通常用 $j=0$ 表示原始信号的尺度，dwt 将原始信号分解为两部分，分别称为逼近系数和细节系数，它们的长度均为原始信号的一半。细节系数就是下一个较粗糙的尺度 $j=1$ 时的离散小波系数。用函数 wavedec 实现离散小波变换时，则可对原始信号进行多级分解。图 1.5 就是用函数 wavedec 实现离散小波变换的结果。图 1.5(a) 为原始信号，信号中混有噪音，而图 1.5(b)、(c)、(d) 分别表示 $j=1、2、3$ 时的离散小波系数。横轴为时间参数 n，纵轴为信号幅度或离散小波系数的幅度。由图 1.5 可以看到，在信号幅值发生突变时，离散小波系数的模较大，这就生动地说明了确实可以用小波变换实现信号的奇异性检测。如果选取一个适当的阈值，只保留模大于该阈值的小波系数，而将其他小波系数的值置零，就实现了信号的数据压缩和去噪。

1.5　常用小波函数

小波分析中用到的小波函数有很多种，具有不惟一性。但小波分析在工程应用中，一个十分重要的问题是最优小波基的选择。

以下介绍几个描述小波特性的基本概念：

1）支撑宽度：小波函数在时域是局部化的，即当 $t \to \pm\infty$ 时，$\psi(t)$ 速降至零。如果一个小波在某个有限持续时间之内是衰减振荡的，而在此持续时间外为零，则称这样的小波是紧支撑的，持续时间的长度就称为该小波的支撑宽度。例如哈尔小波和 Daubechies 小波就是紧支撑

的。而某些小波仅当 $t \to \pm\infty$ 时，才趋于零，这些小波就不是紧支撑的，例如墨西哥帽小波和 Meyer 小波就是这样的小波。对这样的小波就只能定义其等效支撑宽度。

2）正交性：有些小波经过二进伸缩和平移构成 $L^2(R)$ 空间的正交基，这些小波被称为正交小波。例如哈尔小波和 Daubechies 小波、symlets 小波、coiflets 小波都是正交小波，而且它们都是紧支撑的。但并不是所有的小波经过二进伸缩和平移都能构成 $L^2(R)$ 空间的正交基，这些小波就是非正交的。例如墨西哥帽小波和高斯小波都是非正交小波。

3）对称性：如果小波函数关于某一点对称或反对称，则具有对称性。Daubechies 已经证明除哈尔小波外，所有紧支撑的正交小波都不可能是对称的。墨西哥帽小波、高斯小波和 Meyer 小波是对称的，而 Daubechies 小波则是非对称性的。

4）正则性（Regularity）：正则性是用来描述函数光滑程度的。一个时间函数越光滑，当 $\omega \to \pm\infty$ 时，其傅立叶变换衰减越快。所以时间函数的整体正则性可由其傅立叶变换衰减的快慢程度来度量。如 $f(t)$ 的傅立叶变换满足

$$\int_{-\infty}^{+\infty} |\hat{f}(\omega)|(1 + |\omega|)^{1+\alpha} \mathrm{d}\omega < +\infty \quad (\alpha \geqslant 0) \tag{1.5.1}$$

则称 $f(t)$ 正则性为 α，记为 $f \in C^{\alpha}$。当 α 为整数时，$f(t)$ 是 α 次可微的。

例如 Meyer 小波是无限正则的，它无限可微，非常光滑。一般来说，支撑宽度越宽，小波也就越光滑。如 N 阶 Daubechies 小波 dbN 的支撑宽度为 $2N-1$，具有 N 阶消失矩，当 N 较大时，其正则性约为 $0.2N$。

5）消失矩：消失矩是衡量小波性质的一个很重要的参数。小波的 k 阶矩定义为如下积分形式：

$$\int_{-\infty}^{+\infty} t^k \psi(t) \mathrm{d}t$$

如小波的 0 阶、1 阶、直到 $K-1$ 阶矩为 0，即

$$\int_{-\infty}^{+\infty} t^k \psi(t) \mathrm{d}t = 0, \quad k = 0, 1, \cdots, K-1 \tag{1.5.2}$$

则称小波 $\psi(t)$ 具有 K 阶消失矩。由傅立叶变换的性质，上式在频率域的形式为

$$\left.\frac{\mathrm{d}^k \hat{\psi}(\omega)}{\mathrm{d}\omega^k}\right|_{\omega=0} = 0, \quad k = 0, 1, \cdots, K-1 \tag{1.5.3}$$

一般来说，比较高的消失矩意味着小波在时域内振荡次数的增加及频域内更好的局部性。

对很多类型的信号来说，在一定局部时间范围内都能用适当次数的多项式很好地逼近。设信号 f 在区间 $(t_0 - h, t_0 + h)$ 上 K 阶可微，则 $K-1$ 次泰勒多项式

$$P(t) = \sum_{k=0}^{k-1} \frac{f^{(k)}(t_0)}{k!}(t - t_0)^k$$

是在区间 $(t_0 - h, t_0 + h)$ 内对 f 很好的逼近，逼近误差

$$R_K = f(t) - P(t) = \frac{f^{(k)}(\xi)}{K!}(t - t_0)^K, \quad \xi \in (t_0 - h, t_0 + h)$$

满足

$$|R_K| \leqslant \frac{|t - t_0|^K}{K!} \sup |f^{(K)}(\xi)|, \quad \xi \in (t_0 - h, t_0 + h)$$

若小波 $\psi(t)$ 具有 K 阶消失矩，信号在 t_0 的连续小波变换为

$$
\begin{aligned}
Wf(a,t_0) &= \sum_{k=0}^{K-1}\frac{f^{(k)}(t_0)}{k!}\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}(t-t_0)^k\psi\left(\frac{t-t_0}{a}\right)\mathrm{d}t+\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}R_K(t)\psi\left(\frac{t-t_0}{a}\right)\mathrm{d}t\\
&= \sum_{k=0}^{K-1}a^{k+\frac{1}{2}}\frac{f^{(k)}(t_0)}{k!}\int_{-\infty}^{+\infty}x^k\psi(x)\,\mathrm{d}x+\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}R_K(t)\psi\left(\frac{t-t_0}{a}\right)\mathrm{d}t\\
&= \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}R_K(t)\psi\left(\frac{t-t_0}{a}\right)\mathrm{d}t
\end{aligned}
$$

从而有

$$
|Wf(a,t_0)|\leqslant C\int_{-\infty}^{+\infty}x^K\psi(x)\,\mathrm{d}x
$$

可见,小波消失矩越高,小波变换幅值为零或很小的越多,所以消失矩高的小波有利于数据压缩。

下面是几种实际中经常使用的小波:

1)墨西哥帽小波:墨西哥帽小波定义为高斯函数的二阶导数,如图 1.6 所示。这种小波被最先用于计算机视觉中的多尺度边缘检测。规范化的墨西哥帽小波解析表达式是:

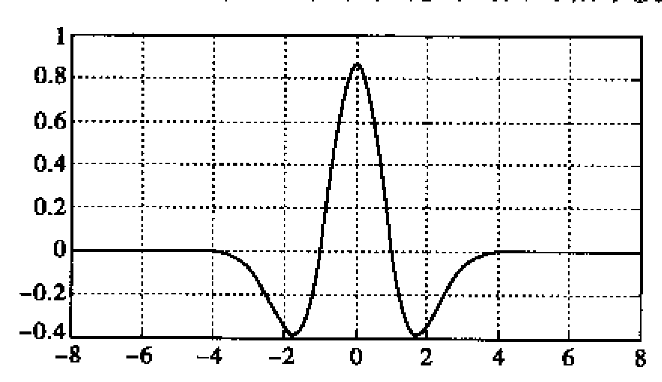


图 1.6 墨西哥帽小波

$$
\psi(t)=\frac{2}{\pi^{1/4}\sqrt{3\sigma}}\left(\frac{t^2}{\sigma^2}-1\right)\exp\left(\frac{-t^2}{2\sigma^2}\right) \tag{1.5.4}
$$

2)复 Morlet 小波:复 Morlet 小波是具有高斯包络的复小波

$$
\psi(t)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{t^2}{2\sigma^2}}e^{i\omega_0 t} \tag{1.5.5}
$$

其傅立叶变换为

$$
\hat{\psi}(\omega)=e^{\frac{\sigma^2(\omega-\omega_0)^2}{2}} \tag{1.5.6}
$$

在信号处理中,如一个函数的傅立叶变换对所有负频率均为 0,则称此函数为解析函数。当 $\omega_0\geqslant 5/\sigma$ 时,对于负频率 $\omega<0$,复 Morlet 小波的傅立叶变换 $\hat{\psi}(\omega)\approx 0$,这时复 Morlet 小波近似为一个解析函数。

3)DOG 小波:DOG 小波为两个不同尺度的高斯函数之差,其解析表达式和傅立叶变换为

$$
\psi(t)=e^{-\frac{t^2}{2}}-\frac{1}{2}e^{-\frac{t^2}{8}} \tag{1.5.7}
$$

$$
\hat{\psi}(\omega)=\sqrt{2\pi}\left(e^{-\frac{\omega^2}{2}}-e^{-2\omega^2}\right) \tag{1.5.8}
$$

4)Daubechies 小波:它是著名数学家 Daubechies 构造的紧支正交小波,它没有具体的解析

表达式。如图 1.7 所示,Daubechies 小波是不对称的。

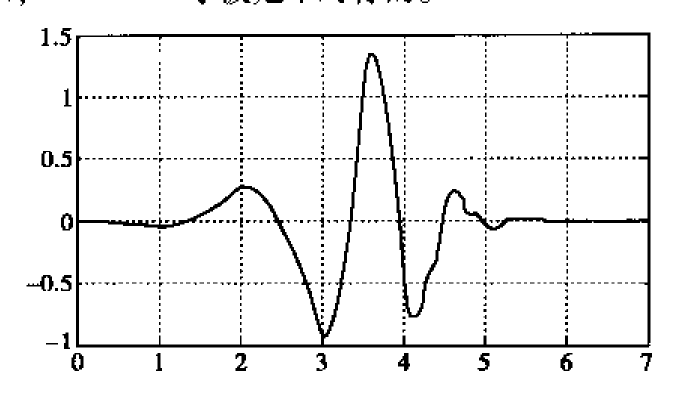


图 1.7 db4 小波

5)Symlets 小波:N 阶 Symlets 小波表示为 sym$N$,它也是由 Daubechies 构造的紧支正交小波,也没有具体的解析表达式。严格说,Symlets 小波也不对称,但与 Daubechies 小波比较,它更接近对称,如图 1.8 所示。

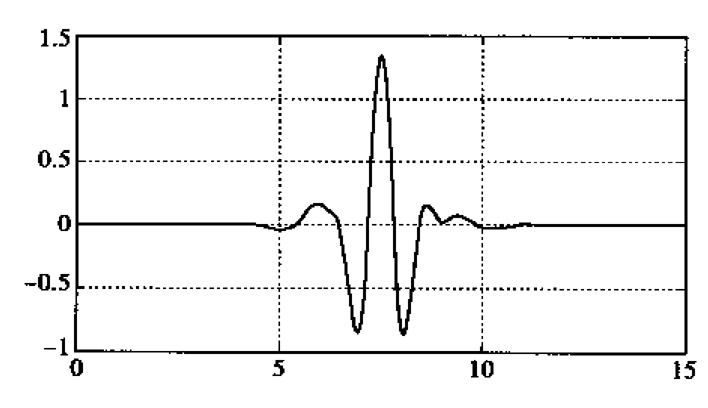


图 1.8 sym8 小波

第 2 章

框 架

框架可以看成正交规范基的推广。在离散小波变换中，正交小波经二进伸缩和平移形成的函数族$\{\psi_{j,n}\}_{j,n\in Z}$构成$L^2(R)$的正交规范基，可以用内积系数$\{\langle f,\psi_{j,n}\rangle\}_{j,n\in Z}$完全描述信号。这时信号在时间域和变换域的能量存在恒等关系，意味着经过变换之后既没丢失信息，也没有信息冗余。一旦放松正交性要求，就意味着放宽能量恒等条件，这就是现在要讨论的框架理论。

在平方可积函数空间中，框架是一族函数$\{\phi_l\}_{l\in J}$，其中l可以是单重也可以是多重指标，指标集J可以是有限也可以是无限的；它由内积系数$\{\phi_l\}_{l\in J}$刻画信号，框架理论研究信号表示的完全性、稳定性与冗余性。在框架理论中，将引出对偶框架$\{\tilde{\phi}_l\}_{l\in J}$，当用框架作信号分解时，可由其对偶框架实现信号重构；也可用对偶框架作信号分解，由框架实现信号重构，即框架和其对偶框架是互为对偶的。由此可见，正交规范基只不过是框架的一种特殊情况，其对偶框架就是它本身，即正交规范基是自对偶的。本章将用框架理论研究信号的二进小波变换。连续小波变换具有平移不变性，但离散小波变换不是平移不变的，这会在模式识别的应用中引起困难。二进小波只对连续小波变换的尺度参数采样，时间参数仍然是连续的，所以保持平移不变性。在计算机视觉中，二进小波变换被用于纹理识别与边缘检测。

2.1 基

在引入框架概念之前，有必要先简要介绍基、正交基和 Riesz 基。

2.1.1 基

定义 2.1 在矢量空间V中，设有一线性无关的矢量集$\{e_l\}_{l\in J}$，属于V的每个矢量f均可表示为e_l的线性组合

$$f = \sum_l c_l e_l, \qquad \forall f \in V \tag{2.1.1}$$

且对任一特定的f，系数集$\{c_l\}_{l\in J}$是惟一的，则称矢量集$\{e_l\}_{l\in J}$为矢量空间V的基，或基底。(2.1.1)式中的l可以是单个指标，也可以是多个指标的集合；J是任意维指标集。

定义 2.2 在 Banach 空间中的基$\{e_l\}_{l\in J}$，如 $\sum_l c_l e_l$ 的收敛仅取决于$|c_l|$，则称为无条件基。

这非常类似于数字序列的绝对收敛。无条件基意味着无穷级数的收敛与各项求和的顺序无关，这在小波分析中有非常重要的意义。对很大一类信号而言，只有少量的小波系数的绝对值很大，也就是说，信号在小波变换域的能量高度集中。这意味着可以用少量幅值较大的小波系数很好地逼近信号，从而实现数据压缩与去噪。

2.1.2 正交基

定义 2.3 在 Hilbert 空间 H 的基底$\{e_l\}_{l\in J}$中，如任意两个不同的矢量正交，即

$$\langle e_l, e_k\rangle = 0, \quad \forall l \neq k$$

则称$\{e_l\}_{l\in J}$为正交基。如正交基中所有矢量的范数均为 1，即

$$\|e_l\| = 1, \quad l \in J$$

则称其为正交规范基。

对正交规范基，其正交性与规范性表示为

$$\langle e_l, e_k\rangle = \delta_{l,k} \tag{2.1.2}$$

对线性组合表达式(2.1.1)两边求内积可得线性组合表达式的系数

$$c_l = \langle f, e_l\rangle \tag{2.1.3}$$

也就是说，对正交规范基而言，线性组合表达式的系数 c_l 很容易计算，它等于待展开的矢量 f 与相应基矢量 e_l 的内积，这是正交规范基的一大优点。将(2.1.3)代入(2.1.1)可得

$$f = \sum_l \langle f, e_l\rangle e_l, \quad \forall f \in H \tag{2.1.4}$$

可以看到，对矢量进行分解和重构都用的是同一正交规范基，这是正交规范基的另一特点。将(2.1.4)式两边对 f 求内积得

$$\|f\|^2 = \sum_l |\langle f, e_l\rangle|^2 = \sum_l |c_l|^2 \tag{2.1.5}$$

这就是正交展开的能量恒等式，它是正交规范基的又一特点。

例 2.1 三维欧几里德空间 R^3 的正交规范基。

在三维欧几里德空间 R^3 中，两个矢量的内积就是它们的点乘积。沿 3 个相互垂直的坐标轴 x,y,z 取 3 个相互垂直的单位长度矢量 $\vec{i},\vec{j},\vec{k}$，它们构成 R^3 的正交规范基。从坐标原点到三维空间中任一点的矢量 $\vec{r}$ 均可表示为

$$\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$$

其中，内积系数

$$x = \vec{r}\cdot\vec{i}, y = \vec{r}\cdot\vec{j}, z = \vec{r}\cdot\vec{k}$$

正是该点的坐标值。可以看到，展开式的系数很容易计算，它们是被展开的矢量 $\vec{r}$ 与对应基矢量 $\vec{i},\vec{j},\vec{k}$ 的内积。

例 2.2 傅立叶级数。

正交规范基在信号处理和图像处理的各种正交变换中获得了广泛的应用，例如众所周知的傅立叶级数。对周期为 T_0 的任意周期函数 $f(t)\in L^2[0,T_0]$，函数族 $\left\{ e^{in\omega_0 t} / \sqrt{T_0} \right\}_{n\in Z}$构成

正交规范基，其中 $\omega_0=2\pi/T_0$ 为基波频率。显然

$$\frac{1}{T_0}\langle e^{in\omega_0 t},e^{im\omega_0 t}\rangle=\frac{1}{T_0}\int_0^{T_0}e^{i(n-m)\omega_0 t}dt=\delta_{n,m}$$

也就是说，不同频率的各次谐波是相互正交的。这时周期信号的线性组合表达式就是傅立叶级数

$$f(t)=\sum_{n=-\infty}^{+\infty}c_n e^{in\omega_0 t}$$

其中的内积系数正是傅立叶系数

$$c_n=\frac{1}{T_0}\langle f,e^{in\omega_0 t}\rangle=\frac{1}{T_0}\int_0^{T_0}f(t)e^{-in\omega_0 t}dt$$

它度量了信号各次谐波分量的幅度和相位信息。应再一次注意到，展开式的系数很容易计算，它们是被展开的矢量 $f(t)$ 与对应基矢量的内积。

例 2.3 离散小波变换。

在离散小波变换中，正交小波经二进伸缩和平移形成的函数族 $\{\psi_{j,n}\}_{j,n\in Z}$ 是 $L^2(R)$ 的正交规范基，内积系数 $\{\langle f,\psi_{j,n}\rangle\}_{j,n\in Z}$ 可以理解为连续小波变换在时间-尺度平面的离散网格上的采样，它描述了信号在这些离散网格点附近的时频窗内的信息。

在小波多分辨率分析中，将引入尺度函数。和小波函数类似，尺度函数经二进伸缩和平移构成它线性张成的函数空间的正交规范基。

定义 2.4 设 $\phi(t)\in L^2(R)$，如通过整数平移形成的函数族 $\{\phi(t-n)\}_{n\in Z}$ 满足

$$\langle\phi(t-k),\phi(t-l)\rangle=\delta_{k,l},\quad k,l\in Z \tag{2.1.6}$$

则称 $\{\phi(t-n)\}_{n\in Z}$ 是一个正交规范基。

下面的定理给出了在频域内某一函数整数平移形成的函数族构成正交规范基的充分必要条件。

定理 2.1 $\phi(t)\in L^2(R)$ 通过整数平移形成的函数族 $\{\phi(t-n)\}_{n\in Z}$ 构成正交规范基，当且仅当

$$\sum_{k=-\infty}^{+\infty}|\hat{\phi}(\omega+2\pi k)|^2=1,\quad\forall\omega\in R \tag{2.1.7}$$

或

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}|\hat{\phi}(\omega)|^2e^{i\omega l}d\omega=\delta_l,\quad l\in Z \tag{2.1.8}$$

证明 事实上，(2.1.6)式等价于

$$\langle\phi(t),\phi(t-l)\rangle=\delta_l,\quad l\in Z \tag{2.1.9}$$

根据 *Parseval* 定理，有

$$\langle\phi(t),\phi(t-l)\rangle=\int_{-\infty}^{+\infty}\phi(t)\phi^*(t-l)dt=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{\phi}(\omega)\hat{\phi}^*(\omega)e^{i\omega l}d\omega$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}|\hat{\phi}(\omega)|^2e^{i\omega l}d\omega=\delta_l$$

(2.1.8)式得证。令上式中 $l=0$，又有

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}|\hat{\phi}(\omega)|^2d\omega=\frac{1}{2\pi}\sum_{k=-\infty}^{+\infty}\int_{(2k-1)\pi}^{(2k+1)\pi}|\hat{\phi}(\omega)|^2d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}\omega \sum_{k=-\infty}^{+\infty} |\hat{\phi}(\omega + 2\pi k)|^2 = 1$$

从而,有 $\sum_{k=-\infty}^{+\infty} |\hat{\phi}(\omega + 2\pi k)|^2 = 1$,这正是(2.1.7)式。

2.1.3 Riesz 基

在信号处理中,选择了一组基,就意味着用系数集去刻画信号的特征。在傅立叶级数中,傅立叶系数度量了信号各次谐波分量的幅度和相位信息。在离散小波变换中,小波系数给出了信号在时间尺度平面上离散网格点附近的时频信息。正交规范基对信号的刻画没有信息冗余,在信号处理和图像处理的各种正交变换中获得了广泛的应用。如果放松对正交性的要求,就可以用更多的方式去描述信号。这正是本章要讨论的框架。在讨论框架之前,先介绍在小波分析中很重要的 Riesz 基。

在无穷维空间中,如果削弱正交性的要求,那么必须加上一个部分的能量等价式以保证基的稳定性。

定义 2.5 称矢量集 $\{e_l\}_{l\in J}$ 是 Hilbert 空间 H 的一个 Riesz 基,如果它是线性无关的,且存在 $0<A$ 和 $B<+\infty$ 使得对任意的 $f\in H$,总可找到 $\{c_l\}_{l\in J}$ 满足

$$f = \sum_l c_l e_l, \quad \forall f \in H \tag{2.1.10}$$

且

$$A\|f\|^2 \leqslant \sum_l |c_l|^2 \leqslant B\|f\|^2 \tag{2.1.11}$$

当 $A=B=1$,且 $\|e_l\|=1$ 时,Riesz 基就是正交规范基。可见 Riesz 基的条件比正交规范基弱。一般来说,(2.1.10)式中的系数 $c_l\neq\langle f,e_l\rangle$。

定理 2.2 $\theta(t)\subset L^2(R)$ 通过整数平移形成的函数族 $\{\theta(t-n)\}_{n\in Z}$ 是它生成的空间的 Riesz 基,当且仅当存在 $0<A$ 和 $B<+\infty$,使得

$$\frac{1}{B} \leqslant \sum_{k=-\infty}^{+\infty} |\hat{\theta}(\omega + 2\pi k)|^2 \leqslant \frac{1}{A}, \quad \forall \omega \in R \tag{2.1.12}$$

证明 由于 $\{\theta(t-n)\}_{n\in Z}$ 是它生成的空间的 Riesz 基,则属于该空间的任意函数 $f(t)$ 可分解成

$$f(t) = \sum_{n=-\infty}^{+\infty} c(n)\theta(t-n)$$

对上式两边做傅立叶变换,得

$$\hat{f}(\omega) = C(\omega)\hat{\theta}(\omega)$$

其中,$C(\omega)$ 是系数序列 $c(n)$ 的傅立叶变换,它是 2π 的周期函数,且满足如下 Plancherel 定理

$$\sum_{n=-\infty}^{+\infty} |c(n)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |C(\omega)|^2 \mathrm{d}\omega \tag{2.1.13}$$

应用 Plancherel 定理,$f(t)$ 的范数可写成

$$\|f\|^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |\hat{f}(\omega)|^2 \mathrm{d}\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |C(\omega)|^2 |\hat{\theta}(\omega)|^2 \mathrm{d}\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}\omega |C(\omega)|^2 \sum_{k=-\infty}^{+\infty} |\hat{\theta}(\omega + 2\pi k)|^2$$

如(2.1.12)式成立,则由上式可得

$$A\|f\|^2 \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi}|C(\omega)|^2 d\omega \leqslant B\|f\|^2$$

将(2.1.13)式代入上式,可得

$$A\|f\|^2 \leqslant \sum_{n=-\infty}^{+\infty}|c(n)|^2 \leqslant B\|f\|^2$$

这就证明了$\{\theta(t-n)\}_{n\in Z}$确实是它生成的空间的 Riesz 基。

2.1.4 正交化

定理 2.3 如$\theta(t)\in L^2(R)$通过整数平移形成的函数族$\{\theta(t-k)\}_{k\in Z}$不是正交规范基,则可以通过

$$\hat{\phi}(\omega) = \frac{\hat{\theta}(\omega)}{\left[\sum_l |\hat{\theta}(\omega+2\pi l)|^2\right]^{\frac{1}{2}}} \tag{2.1.14}$$

得到$\phi(t)\in L^2(R)$,它整数平移形成的函数族$\{\phi(t-k)\}_{k\in Z}$是正交规范基。

实际上,由上式很容易得到

$$\sum_k |\hat{\phi}(\omega+2\pi k)|^2 = \frac{\sum_k |\hat{\theta}(\omega+2\pi k)|^2}{\sum_{l'} |\hat{\theta}(\omega+2\pi l')|^2} = 1$$

由定理 2.1 知,$\phi(t)$整数平移形成的函数族确实是正交规范基。

2.2 框 架

如前所述,若函数族$\{\phi_l(t)\}_{l\in J}$是$L^2(R)$的正交规范基,则任意信号$f(t)\in L^2(R)$均可表示为正交规范基的线性组合

$$f(t) = \sum_{l\in J} \langle f,\phi_l \rangle \phi_l(t) \tag{2.2.1}$$

也就是说,可以用内积系数集$\{\langle f,\phi_l\rangle\}_{l\in J}$去刻画信号。而且用内积系数集去描述信号特征时,没有丢失信号的信息,因为可以用内积系数集按(2.2.1)式重构信号。这称之为正交规范基的完全性。由于正交规范基是线性无关的,所以不存在信息冗余。

框架可以理解为正交规范基的推广。具体地说,框架是 Hilbert 空间 H 的一个矢量族$\{\phi_l\}_{l\in J}$,但一般来说,它是线性相关的。H 可以是平方可积的函数空间 $L^2(R)$,也可以是其他类型的 Hilbert 空间。但在小波分析中,最感兴趣的当然是 H 为 $L^2(R)$的情况。框架是用内积系数集$\{\langle f,\phi_l\rangle\}_{l\in J}$去表征$f$,所以框架理论涉及如下问题:

1)完全性:用内积系数集去表征f时,不应丢失信息。实际上,当框架线性相关时,存在信息冗余,所谓的框架界常数度量了信息冗余的程度。

2)重构:在具有完全性的前提下,可以由内积系数重构f。框架理论引入了对偶框架$\{\tilde{\phi}_l\}_{l\in J}$,任意$f\in H$均可表示为

$$f = \sum_{l \in J} \langle f, \phi_l \rangle \tilde{\phi}_l \tag{2.2.2}$$

也就是说，可以用框架对f进行分解，而用其对偶框架实现重构。而且框架及其对偶框架是互为对偶的，即任意$f \in H$也可表示为

$$f = \sum_{l \in J} \langle f, \tilde{\phi}_l \rangle \phi_l \tag{2.2.3}$$

3）稳定性：可以找到一个数值稳定的算法，由内积系数去计算f。这就要求内积系数有一定误差（如量化误差）时，重构的f不至偏离太大。

这三个问题实际上是相互关联的。下面从框架的定义着手来讨论这些问题。

2.2.1 框架的定义及框架算子

定义2.6 在Hilbert空间中的一族矢量$\{\phi_l\}_{l \in J}$称为一个框架，若存在两个正常数A和B，$0<A$和$B<+\infty$，使得$\{\phi_l\}_{l \in J}$满足

$$A\|f\|^2 \leqslant \sum_{l \in J} |\langle f, \phi_l \rangle|^2 \leqslant B\|f\|^2, \forall f \in H \tag{2.2.4}$$

其中A，B称为框架界。

若$A = B$，称之为紧框架。若$A = B = 1$，且$\|\phi_l\|^2 = 1$，那么框架就是正交规范基。这一结论很容易证明，令(2.2.4)式中$f = \phi_k$，则有

$$\|\phi_k\|^2 = \sum_{l \in J} |\langle \phi_k, \phi_l \rangle|^2 = \|\phi_l\|^4 + \sum_{\substack{l \in J \\ l \neq k}} |\langle \phi_k, \phi_l \rangle|^2$$

由于$\|\phi_k\|^2 = 1$，$\|\phi_l\|^4 = 1$，故上式中$\sum_{\substack{l \in J \\ l \neq k}} |\langle \phi_k, \phi_l \rangle|^2 = 0$。这意味着：当$k \neq l$时，$\langle \phi_k, \phi_l \rangle = 0$。即构成框架的矢量是两两正交的。可见正交规范基是框架的一种特殊情况。

内积系数是一个数字序列，故(2.2.4)式中$\sum_{l \in J} |\langle f, \phi_l \rangle|^2$表示该数字序列的范数的平方。所以，框架定义了一种能量约束条件，可以从两方面来理解它：

1）正变换的连续性：如f与g很接近，则$\{\langle f, \phi_l \rangle\}_{l \in J}$也必定与$\{\langle g, \phi_l \rangle\}_{l \in J}$很接近。由(2.2.4)式，有

$$\sum_{l \in J} |\langle f, \phi_l \rangle - \langle g, \phi_l \rangle|^2 \leqslant B\|f - g\|^2$$

当f无限接近g时，上式右边趋近于0，从而$\{\langle f, \phi_l \rangle\}_{l \in J}$无限接近$\{\langle g, \phi_l \rangle\}_{l \in J}$。

2）重构的连续性：当$\{\langle f, \phi_l \rangle\}_{l \in J}$与$\{\langle g, \phi_l \rangle\}_{l \in J}$十分接近时，$f$与$g$也十分接近。由(2.2.4)式，有

$$A\|f - g\|^2 \leqslant \sum_{l \in J} |\langle f, \phi_l \rangle - \langle g, \phi_l \rangle|^2$$

当$\{\langle f, \phi_l \rangle\}_{l \in J}$无限接近$\{\langle g, \phi_l \rangle\}_{l \in J}$时，上式右边趋近于0，从而$f$无限接近$g$。重构的连续性实际上保证了重构数值算法的稳定性。

下面继续讨论框架理论时，将引入算子这一抽象的数学描述。简单地说，算子实际上是一类运算规则的符号简化表示，它体现了运算规则的共性，但模糊了具体运算类别包含的物理意义。在小波分析中，Hilbert空间H为平方可积的函数空间$L^2(R)$，所以在解释某些抽象的数学描述时，常常以此为例，以便理解。

首先引入对称算子的定义：对任意的$f \in H$，若$T_1 f$和$T_2 f$仍属于H，则称T_1和T_2为H上的

对称算子。当 H 为 $L^2(R)$ 时，算子 T_1 和 T_2 均表示对函数 $f(t)$ 进行某种数学运算，其运算结果 $T_1 f$ 和 $T_2 f$ 仍然是函数。若 $\langle T_1 f,f\rangle \leqslant \langle T_2 f,f\rangle$，则不等式可表示为 $T_1 \leqslant T_2$。算子比较大小本来是很抽象的，但将不等式 $T_1 \leqslant T_2$ 的意义理解为 $\langle T_1 f,f\rangle \leqslant \langle T_2 f,f\rangle$，就比较具体了。

由框架的定义可以看到，框架实际上定义了一个从 H 到 l^2 的映射，框架算子 T 定义为：$(Tf)_l = \langle f,\phi_l\rangle$，即将任意的 $f \in H$ 映射为一个平方可和的序列 $\{\langle f,\phi_l\rangle\}_{l\in J}$，表示为

$$T: L^2(R) \to l^2(J)$$

如 T^* 为 T 的伴随算子，则对任意的 $f\in H$ 和任意的 $C = \{c_l\}_{l\in J} \in l^2(J)$，有

$$\langle f, T^* C\rangle = \langle Tf, C\rangle \tag{2.2.5}$$

这意味着 $T^* C = \sum_{l\in J} c_l \phi_l \in H$，用 $T^*: l^2(J) \to H$ 记之。当 H 为 $L^2(R)$ 时，T^* 将任意的平方可和序列映射为一个平方可积的函数。

显然，$T^* T$ 是一个 H 上的对称算子。因为对任意的 $f\in H$，有

$$T^* Tf = T^* \{\langle f,\phi_l\rangle\}_{l\in J} = \sum_{l\in J} \langle f,\phi_l\rangle \phi_l \tag{2.2.6}$$

将上式两边与 f 求内积可得

$$\langle T^* Tf, f\rangle = \sum_{l\in J} \langle f,\phi_l\rangle \langle f,\phi_l\rangle^* = \sum_{l\in J} |\langle f,\phi_l\rangle|^2 \tag{2.2.7}$$

这就意味着定义框架的不等式(2.2.4)式可以写为

$$AI_d \leqslant T^* T \leqslant BI_d \tag{2.2.8}$$

其中，I_d 为恒等算子。上式说明，对称算子 $T^* T$ 是正有界的线性算子，且有一个严格正常数下界。故可定义其逆算子 $(T^* T)^{-1}$，它也是 H 上的对称算子且满足如下不等式

$$B^{-1} I_d \leqslant (T^* T)^{-1} \leqslant A^{-1} I_d \tag{2.2.9}$$

可以验证：如一有界算子 S 具有一有界逆算子 S^{-1}，且 S 是自伴算子，即 $S^* = S$，那么 $(S^{-1})^* = S^{-1}$，即逆算子也是自伴算子。根据框架算子及其伴随算子的定义很容易证明：$<f, T^* Tg> = <T^* Tf, g>$，即 $(T^* T)^* = T^* T$，也就是说 $T^* T$ 是自伴算子。于是，$(T^* T)^{-1}$ 也是自伴算子，即 $\langle f, (T^* T)^{-1} g\rangle = \langle (T^* T)^{-1} f, g\rangle$。

2.2.2 对偶框架

定义 2.7 定义

$$\tilde{\phi}_l = (T^* T)^{-1} \phi_l \tag{2.2.10}$$

那么，$\{\tilde{\phi}_l\}_{l\in J}$ 构成另一个框架，称之为 $\{\phi_l\}_{l\in J}$ 的对偶框架。

关于框架及其对偶框架，有下述定理。

定理 2.4

1) $\{\tilde{\phi}_l\}_{l\in J}$ 构成另一个框架，且其框架界为 B^{-1} 和 A^{-1}，即

$$B^{-1}\|f\|^2 \leqslant \sum_{l\in J} |\langle f, \tilde{\phi}_l\rangle|^2 \leqslant A^{-1}\|f\|^2, \forall f \in H \tag{2.2.11}$$

对偶框架算子 $\tilde{T}$ 为

$$\tilde{T} = T(T^* T)^{-1} \tag{2.2.12}$$

且对偶框架算子 $\tilde{T}$ 及其伴随算子 $\tilde{T}^*$ 满足

$$\tilde{T}^*\tilde{T} = (T^*T)^{-1} \tag{2.2.13}$$

$$\tilde{T}^* = (T^*T)^{-1}T^* \tag{2.2.14}$$

2）框架算子、对偶框架算子及它们的伴随算子满足

$$\tilde{T}^*T = T^*\tilde{T} = I_d \tag{2.2.15}$$

3）$\tilde{T}T^* = T\tilde{T}^*$ 是 $l^2(J)$ 到 T 的值域的正交投影算子

证明

1）对任意的 $f\in H$，由 $\tilde{\phi}_l$ 的定义，有

$$\begin{aligned}(\tilde{T}f)_l &= \langle f,\tilde{\phi}_l\rangle = \langle f,(T^*T)^{-1}\phi_l\rangle \\ &= \langle (T^*T)^{-1}f,\phi_l\rangle \\ &= (T(T^*T)^{-1}f)_l\end{aligned}$$

于是，(2.2.12)式得证。

上面已得到

$$\langle f,\tilde{\phi}_l\rangle = \langle (T^*T)^{-1}f,\phi_l\rangle$$

其中 $(T^*T)^{-1}f$ 仍然属于 H，参照(2.2.7)式，由上式可得

$$\begin{aligned}\sum_l |<f,\tilde{\phi}_l>|^2 &= \langle \tilde{T}^*\tilde{T}f,f\rangle \\ &= \langle T^*T(T^*T)^{-1}f,(T^*T)^{-1}f\rangle \\ &= \langle f,(T^*T)^{-1}f\rangle \\ &= \langle (T^*T)^{-1}f,f\rangle\end{aligned}$$

于是，(2.2.13)式得证。且由(2.2.9)式推知，(2.2.11)式成立。由(2.2.13)式，有

$$\tilde{T}^*\tilde{T} = (T^*T)^{-1}T^*T(T^*T)^{-1}$$

再从(2.2.12)式可得(2.2.14)式。

2）将(2.2.12)式代入(2.2.13)式，有 $\tilde{T}^*T(T^*T)^{-1} = (T^*T)^{-1}$，所以 $\tilde{T}^*T = I_d$。而由(2.2.12)式，$T^*\tilde{T} = T^*T(T^*T)^{-1} = I_d$。于是，(2.2.15)式得证。

3）由(2.2.12)式和(2.2.14)式易得，$\tilde{T}T^* = T(T^*T)^{-1}T^* = T\tilde{T}^*$。

T 的值域或像集是所有 $f\in H$ 的系数序列 $\{\langle f,\phi_l\rangle\}_{l\in J}$ 的集合，记为 $\mathrm{Im}(T)$。由框架的定义知，$\mathrm{Im}(T)\subset l^2(J)$。但并非所有的属于 $l^2(J)$ 的序列都属于 $\mathrm{Im}(T)$，即 $\mathrm{Im}(T)\neq l^2(J)$。假定框架 $\{\phi_l\}_{l\in J}$ 线性相关，则存在系数不全为零的 $C=\{c_l\}_{l\in J}\in l^2(J)$，使得 $\sum\limits_{l\in J} c_l\phi_l = 0$。从而对任意 $f\in H$，有 $\sum\limits_{l\in J} c_l <f,\phi_l>^* = 0$，即 $C\perp \mathrm{Im}(T)$，这意味着存在属于 $l^2(J)$ 但却与 $\mathrm{Im}(T)$ 正交的序列，所以 $\mathrm{Im}(T)\neq l^2(J)$。用 $\mathrm{Im}^\perp(T)$ 表示 $\mathrm{Im}(T)$ 在 $l^2(J)$ 中的正交补，则 $C\perp \mathrm{Im}(T)$ 意味着 $C\in \mathrm{Im}^\perp(T)$。为证明 $\tilde{T}T^*$ 是 $l^2(J)$ 到 T 的值域的正交投影算子，从两方面加以证明：设 C 为平方可和序列，如 $C\in \mathrm{Im}(T)$，则总存在 f 使 $C=Tf$，从而有 $\tilde{T}T^*C = T(T^*T)^{-1}T^*Tf = Tf = C$，即属于 $\mathrm{Im}(T)$ 的序列的投影就是序列本身；另一方面，如 $C\in \mathrm{Im}^\perp(T)$，则对任意的 $f\neq 0$ 有

$\sum_{l \in J} c_l <f,\phi_l>^* = <\sum_{l \in J} c_l\phi_l, f> = 0$,因为$f \neq 0$,必然是 $\sum_{l \in J} c_l\phi_l = T^* C = 0$,进而 $\tilde{T} T^* C = 0$,即属于 $\mathrm{Im}^{\perp}(T)$ 的序列的投影等于零。

2.2.3 重构[1]

为什么要引入对偶框架呢?由(2.2.15)式 $\tilde{T}^* T = I_d$ 可得 $\tilde{T}^* Tf = f$,从而

$$f = \tilde{T}^* Tf = \tilde{T}^* \{\langle f,\phi_l \rangle\}_{l \in J} = \sum_{l \in J} \langle f,\phi_l \rangle \tilde{\phi}_l$$

即

$$f = \sum_{l \in J} \langle f,\phi_l \rangle \tilde{\phi}_l \tag{2.2.16}$$

上式说明:如$\{\phi_l\}_{l \in J}$构成一个框架,那么任意$f \in H$都可由它的内积系数$\{\langle f,\phi_l \rangle\}_{l \in J}$充分描述,因为可以按上式由内积系数去重构$f$。

同理,由 $\tilde{T}^* T = I_d$,可得

$$f = \sum_{l \in J} \langle f,\tilde{\phi}_l \rangle \phi_l \tag{2.2.17}$$

这意味着也可以用对偶框架$\{\tilde{\phi}_l\}_{l \in J}$进行分析,而用框架$\{\phi_l\}_{l \in J}$实现重构。这说明对偶关系是相互的,即$\{\phi_l\}_{l \in J}$也是$\{\tilde{\phi}_l\}_{l \in J}$的对偶框架。

由上面讨论可以看到,由内积系数$\{\langle f,\phi_l \rangle\}_{l \in J}$重构$f$的关键是找到 $\tilde{\phi}_l$。下面简要介绍寻找 $\tilde{\phi}_l$ 的途径。为此,将上述算子定义加以简化,再定义如下对称算子

$$M = T^* T, \tilde{M} = \tilde{T}^* \tilde{T} = M^{-1} \tag{2.2.18}$$

M 和 $\tilde{M}$ 都是 H 上的对称算子。具体说来

$$M = \sum_{l \in J} \langle \cdot,\phi_l \rangle \phi_l \tag{2.2.19}$$

由(2.2.8)式可得

$$AI_d \leqslant M \leqslant BI_d \tag{2.2.20}$$

这样一来,$\tilde{\phi}_l$ 的定义(2.2.10)式简化表示成

$$\tilde{\phi}_l = M^{-1}\phi_l \tag{2.2.21}$$

对 $A = B$ 紧框架的情况,$M = AI_d$,$M^{-1} = \tilde{M} = A^{-1}I_d$,从而有

$$\tilde{\phi}_l = A^{-1}\phi_l \tag{2.2.22}$$

$$f = A^{-1} \sum_{l \in J} \langle f,\phi_l \rangle \phi_l \tag{2.2.23}$$

如进一步 $A = B = 1$,那么框架就是正交规范基,则 $\tilde{\phi}_l = \phi_l$ 且

$$f = \sum_{l \in J} \langle f,\phi_l \rangle \phi_l \tag{2.2.24}$$

这就是函数的正交展开式。

对一般情况,可将 M 写为 $2(A+B)^{-1}[I_d - (2(A+B)^{-1}M)]$,从而将 $\tilde{\phi}_l$ 写成如下级数形

式

$$\tilde{\phi}_l = \frac{2}{A+B}\sum_{k=0}^{\infty}\left(I_d - \frac{2}{A+B}M\right)^k \phi_l \tag{2.2.25}$$

这便是计算 $\tilde{\phi}_l$ 的数学表达式。因为

$$\frac{A-B}{A+B}I_d \leqslant I_d - \frac{2}{A+B}M \leqslant \frac{B-A}{B+A}I_d$$

或者说

$$\left\| I_d - \frac{2M}{A+B} \right\| \leqslant \frac{B/A-1}{B/A+1} \tag{2.2.26}$$

所以(2.2.25)式所示的级数总是收敛的，且 B/A 越接近 1，收敛越快。当 $B/A \approx 1$ 时

$$\tilde{\phi}_l \approx \frac{2}{A+B}\phi_l \tag{2.2.27}$$

从而

$$f \approx \frac{2}{A+B}\sum_{l\in J}\langle f,\phi_l\rangle\phi_l \tag{2.2.28}$$

从上面的分析知道，对紧框架及 $B/A \approx 1$，计算 ϕ_l 的对偶 $\tilde{\phi}_l$ 很简单，但对一般情况，则很复杂。

当框架 $\{\phi_l\}_{l\in J}$ 线性相关时，存在系数不全为零的序列 $C = \{c_l\}_{l\in J} \in \mathrm{Im}^{\perp}(T)$，使得 $\sum_{l\in J} c_l\phi_l = 0$，这意味着 $T^* C = 0$。由定理 2.4 知，这也意味着 $\tilde{T}^* C = 0$，即 $\sum_{l\in J} c_l\tilde{\phi}_l = 0$。显然，对任意 $f\in H$，可按下式重构

$$f = \sum_{l\in J}(\langle f,\phi_l\rangle + c_l)\tilde{\phi}_l$$

重构系数的范数为

$$\sum_{l\in J}|\langle f,\phi_l\rangle + c_l|^2 = \sum_{l\in J}|\langle f,\phi_l\rangle|^2 + \sum_{l\in J}|c_l|^2$$

这意味着 f 有无穷多个重构，而(2.2.16)式是重构的“最小解”。或者说，T 有无穷多个左逆算子，$\tilde{T}^*$ 是模有最小上确界的左逆算子。这就保证用 T 进行分解，用 $\tilde{T}^*$ 重构的惟一性。当用 $\tilde{T}$ 进行分解，用 T^* 重构也具有惟一性。这一结论对任意框架都是成立的。

例 2.4　H 为二维欧几里德空间 R^2，$\{\phi_1,\phi_2,\phi_3\}$ 构成一个框架。

$$\phi_1 = e_1, \phi_2 = -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2, \phi_3 = -\frac{1}{2}e_1 - \frac{\sqrt{3}}{2}e_2$$

其中，$e_1 = (1,0)$，$e_2 = (0,1)$ 构成 R^2 的正交规范基。

对任意的 $f\in H$，$f = xe_1 + ye_2$，x 和 y 分别是矢量 f 的横坐标和纵坐标。由于

$$\langle f,\phi_1\rangle = x, \langle f,\phi_2\rangle = -\frac{1}{2}x + \frac{\sqrt{3}}{2}y, \langle f,\phi_3\rangle = -\frac{1}{2}x - \frac{\sqrt{3}}{2}y$$

所以

$$\sum_{l=1}^{3}|\langle f,\phi_l\rangle|^2 = \frac{3}{2}(x^2+y^2) = \frac{3}{2}\|f\|^2$$

故 $\{\phi_1,\phi_2,\phi_3\}$ 构成 H 的一个紧框架，框架界常数 $A = B = \frac{3}{2}$。由于 $\sum_{l=1}^{3}\phi_l = 0$，所以框架线性

相关,框架界常数度量了框架在二维欧几里德空间 R^2 中的冗余度。

这时,$T^*T=\frac{3}{2}I_d$,因而 $\tilde{\phi}_l=(T^*T)^{-1}\phi_l=\frac{2}{3}\phi_l$。(2.2.16)式成为

$$f=\frac{2}{3}\sum_{l=1}^{3}\langle f,\phi_l\rangle\phi_l \tag{2.2.29}$$

这意味着 $\tilde{T}^*=\frac{2}{3}T^*$。对任意非零常数 c,均有 $\sum_{l=1}^{3}c\phi_l=0$,所以任意常数序列 $C=\{c,c,c\}$ 均属于框架算子像集的正交补 $\mathrm{Im}^{\perp}(T)$,且 $\tilde{T}^*C=\frac{2}{3}T^*C=0$。所以,$f$ 还可写成

$$f=\frac{2}{3}\sum_{l=1}^{3}(\langle f,\phi_l\rangle+c)\phi_l$$

但(2.2.29)式是重构的"最小解"。

2.2.4 双正交基[3]

Riesz 基 $\{\phi_l\}_{l\in J}$ 是线性无关的框架,这意味着 $\mathrm{Im}(T)=l^2(J)$。由(2.2.16)式推知 $\{\tilde{\phi}_l\}_{l\in J}$ 也是线性无关的,称之为对偶 Riesz 基。令(2.2.17)式中 $f=\phi_k$,得

$$\phi_k=\sum_{l\in J}\langle\phi_k,\tilde{\phi}_l\rangle\phi_l$$

由上式易看到

$$\langle\phi_k,\tilde{\phi}_l\rangle=\delta_{k,l} \tag{2.2.30}$$

上式意味着 Riesz 基和其对偶基相互正交,称之为双正交基。

令(2.2.11)式中 $f=\phi_k$,得

$$B^{-1}\|\phi_k\|^2\leqslant\sum_{l\in J}|\langle\phi_k,\tilde{\phi}_l\rangle|^2\leqslant A^{-1}\|\phi_k\|^2$$

如基是规范化的,由(2.2.30)式,有

$$A\leqslant 1\leqslant B \tag{2.2.31}$$

由于双正交基是线性无关的,所以没有信息冗余,此时框架界满足(2.2.31)式。

当框架线性相关时,$A>1$,存在信息冗余,A 可解释为最小冗余因子。

2.3 小波框架

正交小波经二进伸缩 $a=2^j$ 和整数平移 $b=2^jn$ 形成的函数族

$$\{\psi_{j,n}(t)=2^{-j/2}\psi(2^{-j}t-n)\}_{j,n\in Z}$$

是 $L^2(R)$ 的正交规范基。

如取尺度参数 $a=a_0^j$,平移参数 $b=a_0^jnb_0$,则任意小波形成的函数族

$$\{\psi_{j,n}(t)=a_0^{-j/2}\psi(a_0^{-j}t-nb_0)\}_{j,n\in Z} \tag{2.3.1}$$

是否构成 $L^2(R)$ 的框架呢？也就是说,是否可以从连续小波变换的离散取值 $\langle f,\psi_{j,n}\rangle$ 重构 $f(t)$ 呢？

关于小波框架,关心两个问题:1)对给定的小波函数,找出一个参数 a_0,b_0 的值域 Rg,当 $(a_0,b_0)\in Rg$ 时,$\{\psi_{j,n}\}_{j,n\in Z}$构成一个框架;2)对 $(a_0,b_0)\in Rg$,计算框架界 A,B 的估值。Daubechies 很详尽地研究了产生小波框架时,ψ、a_0、b_0 必须满足的条件,并估计了相应的框架界[1,2]。下面只简要地引用其结论。

假定 ψ 是实的,规范化的且满足允许条件:

$$C_\psi = \int_0^{+\infty}\frac{|\hat{\psi}(\omega)|^2}{\omega}\mathrm{d}\omega < +\infty$$

定理 2.5 如$\{\psi_{j,n}\}_{j,n\in Z}$构成 $L^2(R)$的一个框架,框架界为 A,B,那么有

$$A\leqslant\frac{C_\psi}{b_0\ln a_0}\leqslant B \tag{2.3.2}$$

与

$$A\leqslant\frac{1}{b_0}\sum_{j=-\infty}^{+\infty}|\hat{\psi}(a_0^j\omega)|^2\leqslant B,\ \forall\,\omega\in R-\{0\} \tag{2.3.3}$$

这是 ψ 通过指数伸缩与整数平移构成框架的必要条件。由上述定理不难看到:不等式(2.3.2)把框架界与采样密度联系起来,(2.3.3)式则意味着频率轴被指数伸缩后的小波所覆盖。

下面是关于小波框架的充分条件。

定理 2.6 定义

$$\beta(\xi) = \sup_{1\leqslant|\omega|\leqslant a_0}\sum_{j=-\infty}^{+\infty}|\hat{\psi}(a_0^j\omega)||\hat{\psi}(a_0^j\omega+\xi)| \tag{2.3.4}$$

$$\Delta = \sum_{\substack{k=-\infty\\k\neq0}}^{+\infty}\left[\beta\left(\frac{2\pi k}{b_0}\right)\beta\left(\frac{-2\pi k}{b_0}\right)\right]^{1/2} \tag{2.3.5}$$

如 a_0 和 b_0 使得

$$A_0 = \frac{1}{b_0}\left(\inf_{1\leqslant|\omega|\leqslant a_0}\sum_{j=-\infty}^{+\infty}|\hat{\psi}(a_0^j\omega)|^2-\Delta\right)>0 \tag{2.3.6}$$

$$B_0 = \frac{1}{b_0}\left(\sup_{1\leqslant|\omega|\leqslant a_0}\sum_{j=-\infty}^{+\infty}|\hat{\psi}(a_0^j\omega)|^2+\Delta\right)<+\infty \tag{2.3.7}$$

则$\{\psi_{j,n}\}_{j,n\in Z}$是 $L^2(R)$的框架。常数 A_0 和 B_0 分别是框架界 A 和 B 的下界和上界。

充分条件(2.3.6)式及(2.3.7)式与必要条件(2.3.3)式类似。如Δ相对于 $\inf_{1\leqslant|\omega|\leqslant a_0}\sum_{j=-\infty}^{+\infty}|\hat{\psi}(a_0^j\omega)|^2$ 较小,则 A_0 和 B_0 接近最优框架界 A 和 B。对固定的伸缩步长 a_0,当平移步长 b_0 减少时,Δ 的值减少。

表 2.1 墨西哥帽函数的框架界

a_0	b_0	A	B	B/A
2	0.25	13.091	14.183	1.083
2	0.5	6.546	7.092	1.083
2	1.0	3.223	3.596	1.116
2	1.5	0.325	4.221	12.986
$2^{1/2}$	0.25	27.273	27.278	1.000 2

续表

a_0	b_0	A	B	B/A
$2^{1/2}$	0.5	13.673	13.639	1.000 2
$2^{1/2}$	1.0	6.768	6.870	1.015
$2^{1/2}$	1.75	0.517	7.276	14.061
$2^{1/4}$	0.25	54.552	54.552	1.000 0
$2^{1/4}$	0.5	27.276	27.276	1.000 0
$2^{1/4}$	1.0	13.586	13.690	1.007
$2^{1/4}$	1.75	2.928	12.659	4.324

例如墨西哥帽函数,它是高斯函数 $e^{-t^2/2}$ 的二阶导数

$$\psi(t)=\frac{2}{\sqrt{3}}\pi^{-1/4}(1-t^2)e^{-\frac{t^2}{2}} \tag{2.3.8}$$

前面的系数是为了使 $\|\psi\|=1$。墨西哥帽是视觉分析上常用的小波。当 $a_0=2,2^{1/2},2^{1/4}$,(2.2.25)式中 k 的上限为 1 时,对各种不同的 b_0 值,其框架界常数如表 2.1 所示。

在构成框架的前提下,如何去计算其对偶框架呢?假定小波函数 ψ 的对偶函数为 $\tilde{\psi}$,$\psi_{j,n}$ 的对偶表示为 $\widetilde{\psi_{j,n}}$,如 $\widetilde{\psi_{j,n}}=\tilde{\psi}_{j,n}$,则意味着对偶框架也是由一个函数 $\tilde{\psi}$ 经伸缩和平移形成,于是对偶框架的计算就比较简单。实际情况是不是这样的呢?为了回答这个问题,引入伸缩算子 D^j 和平移算子 S^n 如下

$$\begin{aligned}(D^jf)(t)&=a_0^{-j/2}f(a_0^{-j}t)\\(S^nf)(t)&=f(t-nb_0)\end{aligned} \tag{2.3.9}$$

显然,$\psi_{j,n}=D^jS^n\psi$。从而由对偶框架的定义(2.2.10)式知,$\psi_{j,n}$ 的对偶为

$$\widetilde{\psi_{j,n}}=(T^*T)^{-1}D^jS^n\psi \tag{2.3.10}$$

其中 T 是框架 $\{\psi_{j,n}|_{j,n\in Z}$ 的框架算子,T^* 是 T 的伴随算子。容易验证,对于所有 $f\in L^2(R)$,

$$T^*TD^jf=D^jT^*Tf \tag{2.3.11}$$

即伸缩算子 D^j 和对称算子 T^*T 是可交换的。故 $(T^*T)^{-1}$ 与 D^j 也是可交换的,从而

$$\widetilde{\psi_{j,n}}=D^m(T^*T)^{-1}S^n\psi=a_0^{-j/2}\widetilde{\psi_{0n}}(a_0^{-j}t) \tag{2.3.12}$$

如果 T^*T 与 S^n 也是可交换的,从而 $(T^*T)^{-1}$ 与 S^n 也可交换,那么可将上式写作

$$\widetilde{\psi_{j,n}}=D^jS^n(T^*T)^{-1}\psi=D^jS^n\tilde{\psi}$$

即可以像生成 $\psi_{j,n}$ 一样,由 $\tilde{\psi}$ 通过伸缩和平移得到 $\widetilde{\psi_{j,n}}$,即 $\widetilde{\psi_{j,n}}=\tilde{\psi}_{j,n}$。但非常不幸的是,$T^*T$ 与 S^n 一般说来是不可交换的,故对偶框架的计算一般来说还是很繁杂的。

2.4 二进小波变换

现在用框架理论来研究二进小波变换及其反变换。在模式识别中,信号描述的平移不变性是十分重要的。当模式平移后,它的数值描述也应当相应的平移,而不是被改变。否则,模

式搜索将特别困难。

连续小波变换是平移不变的。设$f_\tau(t)=f(t-\tau)$,则其连续小波变换为

$$\begin{aligned}Wf_\tau(a,b) &= \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t-\tau)\psi\left(\frac{t-b}{a}\right)dt \\ &= \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)\psi\left(\frac{t-(b-\tau)}{a}\right)dt \\ &= Wf(a,b-\tau)\end{aligned} \tag{2.4.1}$$

也就是说,f_τ 的连续小波变换是f的连续小波变换平移相同时间的结果。

小波框架可以理解为对连续小波变换在离散网格点$(a=a_0^j,b=na_0^jb_0)_{j,n\in Z}$上的采样。可以预见,除非连续平移量 τ 是采样间隔 $a_0^jb_0$ 的整数倍,否则,采样值将没有平移不变性。由(2.4.1)式,f_τ 的连续小波变换的采样值为

$$\langle f_\tau,\psi_{j,n}\rangle = Wf_\tau(a_0^j,na_0^jb_0) = Wf(a_0^j,na_0^jb_0-\tau) \tag{2.4.2}$$

当连续平移量 τ 是采样间隔 $a_0^jb_0$ 的整数倍,即 $\tau=ka_0^jb_0,k\in Z$ 时,

$$\langle f_\tau,\psi_{j,n}\rangle = \langle f,\psi_{j,n-k}\rangle \tag{2.4.3}$$

采样值具有平移不变性。显然,当连续平移量 τ 不是采样间隔 $a_0^jb_0$ 的整数倍时,采样值将没有平移不变性。

二进小波变换只是将尺度参数离散化,时间平移参数仍然是连续的,所以二进小波变换具有平移不变性。有些学者在某些场合使用了卷积型二进小波变换,所以首先介绍卷积型小波变换。

2.4.1 卷积型小波变换的定义和性质

设 $\psi(t)\in L^2(R)$,令

$$\psi_s(t) = \frac{1}{s}\psi\left(\frac{t}{s}\right),s>0 \tag{2.4.4}$$

显然,$\psi_s(t)$的面积与母小波 $\psi(t)$相同。对$f(t)\in L^2(R)$,其卷积型小波变换定义为

$$Wf(s,t) = f*\psi_s(t) = \frac{1}{s}\int_{-\infty}^{+\infty} f(t)\psi\left(\frac{t-\tau}{s}\right)d\tau \tag{2.4.5}$$

与卷积型小波变换相对应,将1.3节定义的小波变换称为相关型小波变换。二者之间是可以相互转换的,令

$$h(t) = \frac{1}{\sqrt{s}}\psi^*(-t)$$

则卷积型小波变换可写成

$$Wf(s,t) = f*\psi_s(t) = \frac{1}{\sqrt{s}}\int_{-\infty}^{+\infty} f(t)h^*\left(\frac{\tau-t}{s}\right)d\tau \tag{2.4.6}$$

也就是说,以$\psi(t)$作为小波函数的卷积型小波变换可看成以$h(t)$作为小波函数的相关型小波变换。卷积型小波变换与相关型小波变换并没有本质区别,所以卷积型小波变换也具有与相关型小波变换相同的性质,如线性、平移不变性、尺度共变性、内积定理等。

2.4.2 二进小波变换[3]

连续小波变换中的尺度参数和平移参数都是连续变化的,离散小波变换则将二者都离散

化了。而二进小波变换只是将尺度参数离散化为2的整次幂，即 $s=2^j, j\in Z$。$\psi_j(t)$ 表示将 $\psi(t)$ 作二进伸缩的同时保持其范数不变，即

$$\psi_j(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t}{2^j}\right),\quad j\in Z \tag{2.4.7}$$

若 $\psi(t)$ 是实小波函数，则 $f(t)\in L^2(R)$ 的二进小波变换定义为

$$Wf(2^j,t) = \int_{-\infty}^{+\infty} f(\tau)\frac{1}{\sqrt{2^j}}\psi\left(\frac{\tau-t}{2^j}\right)\mathrm{d}\tau = f*\bar{\psi}_j(t),\quad j\in Z \tag{2.4.8}$$

其中，$\bar{\psi}_j(t)=\psi_j(-t)$。应该注意，信号 $f(t)$ 的二进小波变换是一个函数序列，其中，j 表示尺度，t 表示时间平移量。

为了能从二进小波变换重构信号，或者说，信号的二进小波变换并未丢失信息，则 $\psi(t)$ 必须满足一定的约束条件。

定理 2.7　若存在两个正数 $0<A\leqslant B<+\infty$，使得下列不等式成立

$$A\leqslant\sum_{j=-\infty}^{\infty}|\hat{\psi}(2^j\omega)|^2\leqslant B,\ \forall\omega\in R-\{0\} \tag{2.4.9}$$

则

$$A\|f\|^2\leqslant\sum_{j=-\infty}^{\infty}\frac{1}{2^j}\|Wf(2^j,t)\|^2\leqslant B\|f\|^2 \tag{2.4.10}$$

如果 $\tilde{\psi}$ 满足

$$\sum_{j=-\infty}^{\infty}\hat{\psi}^*(2^j\omega)\hat{\tilde{\psi}}(2^j\omega) = 1,\ \forall\omega\in R-\{0\} \tag{2.4.11}$$

则

$$f(t) = \sum_{j=-\infty}^{\infty}\frac{1}{2^j}Wf(2^j,t)*\tilde{\psi}_j(t) \tag{2.4.12}$$

其中

$$\tilde{\psi}_j(t) = \frac{1}{\sqrt{2^j}}\tilde{\psi}\left(\frac{t}{2^j}\right),\quad j\in Z \tag{2.4.13}$$

证明　由二进小波变换的定义(2.4.8)式得 $f_j(t)=Wf(2^j,t)$ 的傅立叶变换

$$\hat{f}_j(\omega) = \sqrt{2^j}\hat{f}(\omega)\hat{\psi}^*(2^j\omega) \tag{2.4.14}$$

也就是说，$\hat{f}_j(\omega)$ 相当于信号通过一个带通滤波器，所以二进小波变换相当于信号通过中心频率和带宽不同的带通滤波器组。而(2.4.9)式确保带通滤波器组覆盖整个频率轴，从而不丢失 $f(t)$ 的信息。

由(2.4.9)得

$$A\int_{-\infty}^{+\infty}|\hat{f}(\omega)|^2\mathrm{d}\omega\leqslant\sum_{j=-\infty}^{+\infty}\int_{-\infty}^{+\infty}|\hat{f}(\omega)|^2|\hat{\psi}(2^j\omega)|^2\mathrm{d}\omega\leqslant B\int_{-\infty}^{+\infty}|\hat{f}(\omega)|^2\mathrm{d}\omega$$

将(2.4.14)式代入上式得

$$A\int_{-\infty}^{+\infty}|\hat{f}(\omega)|^2\mathrm{d}\omega\leqslant\sum_{j=-\infty}^{+\infty}\frac{1}{2^j}\int_{-\infty}^{+\infty}|\hat{f}_j(\omega)|^2\mathrm{d}\omega\leqslant B\int_{-\infty}^{+\infty}|\hat{f}(\omega)|^2\mathrm{d}\omega$$

$$A\|\hat{f}(\omega)\|^2\leqslant\sum_{j=-\infty}^{+\infty}\frac{1}{2^j}\|\hat{f}_j(\omega)\|^2\leqslant B\|\hat{f}(\omega)\|^2$$

由 Plancherel 定理：$\|\hat{f}(\omega)\|^2 = 2\pi\|f(t)\|^2$，$\|\hat{f}_j(\omega)\|^2 = 2\pi\|Wf(2^j,t)\|^2$，就可由上式得到(2.4.10)式。

将(2.4.11)式两边同时乘以$\hat{f}(\omega)$并将(2.4.14)式代入得

$$\hat{f}(\omega) = \sum_{j=-\infty}^{+\infty}\hat{f}(\omega)\hat{\psi}^*(2^j\omega)\hat{\tilde{\psi}}(2^j\omega)$$
$$= \sum_{j=-\infty}^{+\infty}\frac{1}{\sqrt{2^j}}\hat{f}_j(\omega)\hat{\tilde{\psi}}(2^j\omega) = \sum_{j=-\infty}^{+\infty}\frac{1}{2^j}\hat{f}_j(\omega)\hat{\tilde{\psi}}_j(\omega)$$

上式取傅立叶反变换便可得到(2.4.12)式。证毕。

(2.4.10)表明$\{Wf(2^j,t)\}_{j\in Z}$构成一个小波框架，因此由它可以重构$f(t)$。B/A 越接近 1，重构的稳定性越好。满足(2.4.11)式的 $\tilde{\psi}$ 称为重构小波，但需要指出的是，满足(2.4.11)式的重构小波将会有无穷多个。如选取

$$\hat{\tilde{\psi}}(\omega) = \frac{\hat{\psi}(\omega)}{\sum_{j=-\infty}^{+\infty}|\hat{\psi}(2^j\omega)|^2} \tag{2.4.15}$$

则重构二进小波 $\tilde{\psi}$ 是惟一的，它相当于框架理论中提到的重构"最小解"。

关于二进小波变换的快速算法及其应用将在第 7 章详细讨论。

第3章 信号的多分辨率分析

小波变换中的伸缩参数实质上描述了观测信号的范围,也就是尺度,在图像处理中,称之为分辨率。所以小波变换也可以理解为信号的多分辨率分析。多分辨率分析和时-频分析一样,是理解小波变换的基本概念。小波多分辨率分析不仅仅是信号分析的一种方法,更重要的是:小波多分辨率分析中的 Mallat 算法其地位和作用类似于快速傅立叶变换算法在傅立叶分析中的地位和作用;在小波多分辨率分析数学模型的推导过程中,将指出构造小波的一条重要途径。此外,本章还将引入许多重要的概念,以后各章内容可以说都是这些概念的进一步发展和深化,所以本章是全书的重点。在小波多分辨率分析中,将引出与小波函数 $\psi(t)$ 紧密相联的另一个分析函数——尺度函数 $\phi(t)$。在 Mallat 算法中,与小波函数和尺度函数相联系的是两个数字滤波器,与 $\phi(t)$ 相联的是低通滤波器 $h(n)$,而与 $\psi(t)$ 相联的是高通滤波器 $g(n)$。Mallat 算法是一个递推计算过程,首先将原始离散信号通过这两个滤波器滤波,再进行下采样(取偶数样本),分别得到较粗分辨率的离散逼近信号和离散细节信号,然后再对离散逼近信号滤波下采样,得到更粗分辨率的离散逼近信号和离散细节信号。而离散细节信号正是离散小波变换系数。

3.1 多分辨率分析

为了避免混淆,我们说尺度等于 2^j 时的分辨率为 2^{-j}。多分辨率分析的基本思想是构造嵌套的线性函数空间序列 $\{V_j\}_{j\in Z}$,j 表示不同的分辨率,V_j 是在分辨率为 2^{-j}时对 $L^2(R)$ 的逼近,分辨率越高,逼近程度越高。具体来说,对任意 $f(t)\in L^2(R)$,其分辨率为 2^{-j}时的逼近信号 $f_{A_j}(t)\in V_j$ 就是 $f(t)$ 在 V_j 上的正交投影,用 A_j 表示正交投影算子,则有

$$A_jf(t)=f_{A_j}(t),f(t)\in L^2(R),f_{A_j}(t)\in V_j,j\in Z \tag{3.1.1}$$

分辨率越高,$f_{A_j}(t)$ 对 $f(t)$ 的逼近程度越高。根据投影定理,A_j 是正交投影算子意味着

$$\| A_jf(t)-f(t)\| \leq \| g(t)-f(t)\|,\forall g(t)\in V_j \tag{3.1.2}$$

上列不等式左边可解释为用 $A_jf(t)$ 逼近 $f(f)$ 的逼近误差,右边为用 $g(t)$ 逼近 $f(t)$ 的逼近误差,故(3.1.2)式意味着 $A_jf(t)$ 是分辨率为 2^{-j}时对 $f(t)$ 的最佳逼近。

定义 3.1　如果满足下面 6 条性质，则 $L^2(R)$ 中的嵌套闭子空间序列 $\{V_j\}_{j\in Z}$ 称为 $L^2(R)$ 的一个多分辨率逼近

$$V_{j+1}\subset V_j,\ \forall j\in Z \tag{3.1.3}$$

$$f(t)\in V_{j+1}\Leftrightarrow f(2t)\in V_j,\ \forall f(t)\in L^2(R),\ \forall j\in Z \tag{3.1.4}$$

$$f(t)\in V_j\Leftrightarrow f(t-2^jk)\in V_j,\ \forall f(t)\in L^2(R),\ \forall j,k\in Z \tag{3.1.5}$$

$$\lim_{j\to+\infty} V_j=\bigcap_{j=-\infty}^{+\infty} V_j=\{0\} \tag{3.1.6}$$

$$\lim_{j\to-\infty} V_j=\text{Closure}\left(\bigcup_{j=-\infty}^{+\infty} V_j\right)=L^2(R) \tag{3.1.7}$$

存在 θ，使得 $\{\theta(t-n)\}_{n\in Z}$ 是 V_0 的 Riesz 基。

现在给出这些数学性质的直观解释。(3.1.3)式是一个因果性质，它表明分辨率 2^{-j} 时的逼近包含为计算较低分辨率 $2^{-(j+1)}$ 时的逼近应有的全部信息，也就是说 $A_jf(t)$ 包含为了计算 $A_{j+1}f(t)$ 所必需的所有信息，或者说 $A_jf(t)$ 比 $A_{j+1}f(t)$ 包含了更多的信息，即 $f_{A_j}(t)$ 比 $f_{A_{j+1}}(t)$ 更好地逼近了 $f(t)$。(3.1.4)式则意味着，如 $f(t)$ 属于 V_{j+1}，则 $f(t)$ 经二进压缩得到的 $f(2t)$ 必然属于 V_j，反之亦然；我们说相邻的两个矢量空间 V_j 和 V_{j+1} 之间具有二尺度关系。分辨率 2^{-j} 意味着在单位时间内对 $A_jf(t)$ 抽取 2^{-j} 个样本，用 $a_j(n)$ 来表示这些离散样本，$A_jf(t)$ 可以被 $a_j(n)$ 完全刻画。(3.1.5)式意味着信号 $f(t)$ 平移了与尺度 2^j 成比例的一段时间 2^jk 成为 $f(t-2^jk)$ 后 $A_jf(t)$ 也平移了同样一段时间，$a_j(n)$ 则平移为 $a_j(n-k)$。(3.1.6)式意味着当 $j\to+\infty$ 时，分辨率越来越低，$A_jf(t)$ 包含越来越少的信息，对 $f(t)$ 逼近得越来越差；而(3.1.7)式意味着当 $j\to-\infty$ 时，分辨率越来越高，$A_jf(t)$ 包含越来越多的信息，对 $f(t)$ 的逼近程度越来越高。

定义 3.2　只要求多分辨率分析存在 Riesz 基，这个条件比较弱，一般要求存在正交基。实际上，可将 Riesz 基 $\{\theta(t-n)\}_{n\in Z}$ 正交化而得到正交基 $\{\phi(t-n)\}_{n\in Z}$，其中 $\phi(t)$ 称为多分辨率分析的尺度函数。

按投影定理，V_j 可分解为 V_{j+1} 及其在 V_j 中的正交补 W_{j+1} 的直和，即

$$V_j=V_{j+1}\oplus W_{j+1},\ V_{j+1}\perp W_{j+1},\ j\in Z \tag{3.1.8}$$

令 D_j 为 $L^2(R)$ 到 W_j 的正交投影算子，则

$$f_{D_j}(t)=D_jf(t) \tag{3.1.9}$$

表示 $f(t)$ 在 W_j 上的正交投影，称为 $f(t)$ 在分辨率为 2^{-j} 时的细节信号。由(3.1.8)式可得

$$f_{A_j}(t)=f_{A_{j+1}}(t)+f_{D_{j+1}}(t) \tag{3.1.10}$$

$f_{A_j}(t)$ 比 $f_{A_{j+1}}(t)$ 包含了更多的信息，而多余的信息包含在 $f_{D_{j+1}}(t)$ 中。稍后将看到，$f_{A_{j+1}}(t)$ 包含了 $f_{A_j}(t)$ 的低频部分，而 $f_{D_{j+1}}(t)$ 包含了 $f_{A_j}(t)$ 的高频部分。

V_j 称为尺度函数空间，而 W_j 称为小波函数空间。多分辨率分析实际上是尺度函数空间不断递推分解的过程。从(3.1.8)式可以看到，如果从 V_L 出发，将 V_L 分解为 V_{L+1} 和 W_{L+1}，又将 V_{L+1} 分解为 V_{L+2} 和 W_{L+2}，再将 V_{L+2} 分解为 V_{L+3} 和 W_{L+3}，如此递推分解下去，直到将 V_{L+J-1} 分解为 V_{L+J} 和 W_{L+J}。于是，多分辨率分析可以表式为

$$V_L=W_{L+1}\oplus W_{L+2}\oplus\cdots\oplus W_{L+j}\oplus\cdots\oplus W_{L+J}\oplus V_{L+J} \tag{3.1.11}$$

将 $f(t)$ 正交投影到上面的函数空间得

$$f_{A_L}(t)=f_{D_{L+1}}(t)+f_{D_{L+2}}+\cdots+f_{D_{L+j}}+\cdots+f_{D_{L+J}}+f_{A_{L+J}} \tag{3.1.12}$$

其中，j 表示分解的层次。一般都假定多分辨率分析从 V_0 开始进行递推分解，这时有

$$V_0 = W_1 \oplus W_2 \oplus \cdots \oplus W_j \oplus \cdots \oplus W_J \oplus V_J \tag{3.1.13}$$

$$f_{A_0}(t) = f_{D_1}(t) + f_{D_2} + \cdots + f_{D_j} + \cdots + f_{D_J} + f_{A_J} \tag{3.1.14}$$

式中 j 既表示分解的层次，同时又表示分辨率为 2^{-j}。

3.2 Mallat 算法

1987 年法国科学家 Stéphan Mallat 将计算机视觉领域内的多分辨率分析思想引入小波分析，推导出相应的快速算法[4]。在多分辨率分析的 Mallat 算法中，将引入另一个重要的分析函数—尺度函数 $\phi(t)$，它经伸缩和平移之后得到函数族 $\{\phi_{j,n}(t)\}_{n\in Z}$ 构成尺度空间 V_j 的正交规范基。于是信号 $f(t)$ 在 V_j 上的正交投影，即逼近信号 $f_{A_j}(t)$ 可表示为在 $\{\phi_{j,n}(t)\}_{n\in Z}$ 上的正交展开式，展开式的系数 $a_j(n)$ 称为离散逼近信号。类似的，正交小波函数 $\psi(t)$ 经二进伸缩和平移之后得到函数族 $\{\psi_{j,n}(t)\}_{n\in Z}$ 构成小波空间 W_j 的正交规范基。$f(t)$ 在 W_j 上的正交投影，既细节信号 $f_{D_j}(t)$ 也可表示为正交展开的形式，展开式的系数 $d_j(n)$ 称为离散细节信号。Mallat 算法正是离散信号在相邻分辨率之间的递推计算方法。在 Mallat 算法中，将不再出现尺度函数和小波函数，而是与它们对应的数字滤波器 $h(n)$ 和 $g(n)$。如原始信号对应着 $j=0$，表示为 $a_0(n)$，它可能是模拟信号的离散采样，也可能本身就是一个离散的数字序列。$a_0(n)$ 经 h 和 g 滤波后再下取样分别得到 $a_1(n)$ 和 $d_1(n)$，$a_1(n)$ 再经这两个数字滤波器滤波和下取样得到 $a_2(n)$ 和 $d_2(n)$，如此递推下去，最后得到 $a_J(n)$ 和 $d_J(n)$。必须注意，离散细节信号 $d_j(n)$ 正是离散小波系数，所以 Mallat 算法也是离散小波变换的快速算法。

3.2.1 分解算法

设 $\{\theta(t-n)\}_{n\in Z}$ 是 V_0 的 Riesz 基。

定理 3.1 设 $\{V_j\}_{j\in Z}$ 为 $L^2(R)$ 的多分辨率逼近，则存在 $\phi(t)\in L^2(R)$，其傅立叶变换为

$$\hat{\phi}(\omega) = \frac{\hat{\theta}(\omega)}{\left(\sum_{k=-\infty}^{+\infty} |\hat{\theta}(\omega+2\pi k)|^2\right)^{1/2}} \tag{3.2.1}$$

且 $\phi(t)$ 经二进伸缩和平移得

$$\phi_{j,n}(t) = 2^{-\frac{j}{2}}\phi(2^{-j}t-n), \quad j,n\in Z$$

则函数族 $\{\phi_{j,n}(t)\}_{n\in Z}$ 为 V_j 的正交规范基

$$\overline{\text{linear span}\{\phi_{j,n}(t)\}_{n\in Z}} = V_j \tag{3.2.2}$$

$\phi(t)$ 称为多分辨逼近 $\{V_j\}_{j\in Z}$ 的尺度函数。

(3.2.2)式意味着 V_j 是由函数族 $\{\phi_{j,n}(t)\}_{n\in Z}$ 张成的线性闭包，即 V_j 中的任意一个函数均可表示为 $\{\phi_{j,n}(t)\}_{n\in Z}$ 的线性组合或其线性组合的极限。$\phi_{j,n}(t)$ 表达式中的 $2^{-\frac{j}{2}}$ 是为了保持其能量与 $\phi(t)$ 的能量一致，即

$$\int_{-\infty}^{+\infty}\phi_{j,n}^2(t)\,dt = \int_{-\infty}^{+\infty}\phi^2(t)\,dt$$

现在开始推导分解算法。由定理 3.1，任意能量有限信号 $f(t)$ 在 V_j 上的正交投影可写为如下正交展开式

$$A_jf(t) = \sum_n \langle f(u), \phi_{j,n}(u) \rangle \phi_{j,n}(t) \tag{3.2.3}$$

上述正交展开式的系数定义为$f(t)$在分辨率为2^{-j}时的离散逼近信号，记为$a_j(n)$，即

$$a_j(n) = \langle f(u), \phi_{j,n}(u) \rangle \tag{3.2.4}$$

离散逼近信号$a_j(n)$又称为尺度系数。

由于$\phi_{j+1,n}(t) \in V_{j+1}$，而$V_{j+1} \subset V_j$，故$\phi_{j+1,n}(t)$可按$V_j$的正交规范基展开为

$$\phi_{j+1,n}(t) = \sum_k \langle \phi_{j+1,n}(u), \phi_{j,k}(u) \rangle \phi_{j,k}(t)$$

上式中，由内积表示的系数为

$$\begin{aligned}\langle \phi_{j+1,n}(u), \phi_{j,k}(u) \rangle &= \int_{-\infty}^{+\infty} 2^{-\frac{j+1}{2}} \phi(2^{-(j+1)}u - n) 2^{-\frac{j}{2}} \phi(2^{-j}u - k) \mathrm{d}u \\ &= 2^{\frac{1}{2}} \int_{-\infty}^{+\infty} \phi(y - n)\phi(2y - k)\mathrm{d}y = 2^{\frac{1}{2}} \int_{-\infty}^{+\infty} \phi(x)\phi(2x + 2n - k)\mathrm{d}x\end{aligned}$$

令

$$h(n) = \sqrt{2} \int_{-\infty}^{+\infty} \phi(t)\phi(2t - n)\mathrm{d}t \tag{3.2.5}$$

则由内积表示的系数可写成

$$\langle \phi_{j+1,n}(u), \phi_{j,k}(u) \rangle = h(k - 2n)$$

从而可得

$$\phi_{j+1,n}(t) = \sum_k h(k - 2n)\phi_{j,k}(t)$$

将上式两边与$f(t)$求内积，有

$$\langle f(t), \phi_{j+1,n}(t) \rangle = \sum_k h(k - 2n) \langle f(t), \phi_{j,k}(t) \rangle$$

注意到离散逼近信号的定义(3.2.4)式，则上式可写为

$$a_{j+1}(n) = \sum_k h(k - 2n) a_j(k) \tag{3.2.6}$$

它正是两相邻分辨率的离散逼近信号之间的递推关系式。

上式中的h是由(3.2.5)式定义的一个数字滤波器，有时称之为尺度滤波器，它在Mallat算法中起核心作用。(3.2.5)式反映了多分辨率分析中两相邻分辨率之间的关系，还可以看到，数字滤波器$h(n)$和尺度函数$\phi(t)$是紧密相连的。

如定义

$$\bar{h}(n) = h(-n) \tag{3.2.7}$$

则(3.2.6)式可更明显地表示为数字滤波的形式

$$a_{j+1}(n) = \sum_k \bar{h}(2n - k) a_j(k) \tag{3.2.8}$$

即$a_j(n)$经过冲激响应为$\bar{h}(n)$的数字滤波器之后再抽取偶数样本就得到$a_{j+1}(n)$。

现在解释抽取偶数样本。令$a'_{j+1}(p)$为$a_j(p)$经过冲激响应为$\bar{h}(p)$的数字滤波器之后的输出，则$a'_{j+1}(p)$可表示为如下卷积和

$$a'_{j+1}(p) = \sum_k \bar{h}(p - k) a_j(k)$$

令$p = 2n$，$a_{j+1}(n) = a'_{j+1}(2n)$意味着$a_{j+1}(n)$是$a'_{j+1}(p)$抽取偶数样本的结果，代入上式便可得(3.2.8)式。

由(3.2.3)和(3.2.4)可将$f(t)$在V_j上的正交投影写为如下正交展开式

$$f_{A_j}(t) = A_j f(t) = \sum_n a_j(n)\phi_{j,n}(t) \tag{3.2.9}$$

它可以看成是由离散逼近信号$a_j(n)$得到连续逼近信号$f_{A_j}(t)$的内插表达式。

关于小波函数空间的正交规范基，有如下定理：

定理 3.2 设$\{V_j\}_{j\in Z}$为多分辨率逼近的矢量空间序列，$\phi(t)$为相应的尺度函数，总存在与它对应的正交小波$\psi(t)$，其傅立叶变换$\hat{\psi}(\omega)$由下式给定

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}G\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right), G(\omega) = e^{-i\omega}H^*(\omega+\pi) \tag{3.2.10}$$

其中$H(\omega)$为$h(n)$的傅立叶变换。由$\psi(t)$二进伸缩及整数平移后得到的函数族

$$\psi_{j,n}(t) = 2^{-\frac{j}{2}}\psi(2^{-j}t-n) \tag{3.2.11}$$

具有如下性质：$\{\psi_{j,n}(t)\}_{n\in Z}$是$W_j$的正交规范基，即

$$\overline{\text{linear span}\{\psi_{j,n}(t)\}_{n\in Z}} = W_j \tag{3.2.12}$$

而$\{\psi_{j,n}(t)\}_{j,n\in Z}$是$L^2(R)$的正交规范基，即

$$\overline{\text{linear span}\{\psi_{j,n}(t)\}_{j,n\in Z}} = L^2(R) \tag{3.2.13}$$

上式意味着，任意$f(t)\in L^2(R)$均可展开为$\{\psi_{j,n}(t)\}_{j,n\in Z}$的线性叠加，这就是第1章讲到的离散小波变换。此外，$G(\omega)$是数字滤波器$g(n)$的傅立叶变换，由(3.2.10)式可导出

$$g(n) = (-1)^{1-n}h(1-n) \tag{3.2.14}$$

由(3.2.10)式还可以看到：从$h(n)$可导出$\psi(t)$。今后将看到，这确实是构造小波的一条重要途径，它在小波分析的发展过程中起着重大的作用。

现在，进一步推导分解算法。根据定理3.2，$f(t)$在W_j上的正交投影为

$$D_j f(t) = \sum_n \langle f(u), \psi_{j,n}(u)\rangle \psi_{j,n}(t) \tag{3.2.15}$$

类似的，将上式中由内积表示的系数定义为分辨率为2^{-j}时的离散细节信号$d_j(n)$，即

$$d_j(n) = \langle f(u), \psi_{j,n}(u)\rangle \tag{3.2.16}$$

显然，离散细节信号$d_j(n)$正是离散小波系数。$f(t)$在W_j上的正交投影可写成

$$f_{D_j}(t) = D_j f(t) = \sum_n d_j(n)\psi_{j,n}(t) \tag{3.2.17}$$

因为$W_{j+1}\subset V_j$，$\psi_{j+1,n}(t)\in W_{j+1}$，所以$\psi_{j+1,n}(t)$可按$V_j$的正交规范基展开为

$$\psi_{j+1,n}(t) = \sum_k \langle \psi_{j+1,n}(u), \phi_{j,k}(u)\rangle \phi_{j,k}(t)$$

经过类似的推导，可得如下递推关系式

$$d_{j+1}(n) = \sum_k g(k-2n)a_j(k) \tag{3.2.18}$$

其中

$$g(n) = \sqrt{2}\int_{-\infty}^{+\infty}\psi(t)\phi(2t-n)\mathrm{d}t \tag{3.2.19}$$

如定义

$$\bar{g}(n) = g(-n) \tag{3.2.20}$$

则(3.2.18)式又可写成数字滤波的形式

$$d_{j+1}(n) = \sum_k \bar{g}(2n-k)a_j(k) \tag{3.2.21}$$

即 $a_j(n)$ 经过数字滤波器 $\bar{g}(n)$ 之后再抽取偶数样本就得到 $d_{j+1}(n)$。

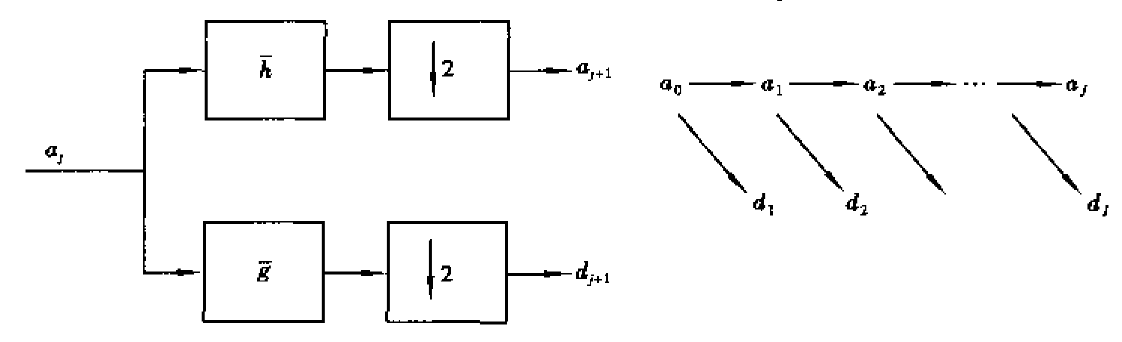


图 3.1　Mallat 信号分解算法

(3.2.8)式和(3.2.21)式就是 Mallat 算法信号分解的表达式，如图 3.1 所示。设原始信号为 $a_L(n)$，分解为 $a_{L+1}(n)$ 和 $d_{L+1(n)}$，$a_{L+1}(n)$ 再分解为 $a_{L+2}(n)$ 和 $d_{L+2}(n)$，进行了 J 次分解之后，得到 $a_{L+J}(n)$ 和 $d_{L+J}(n)$。这样将原始信号 $a_L(n)$ 逐级分解为如下一组离散信号

$$\{a_{L+J};d_{L+j},1\leqslant j\leqslant J\} \tag{3.2.22}$$

其中 j 表示分解的级次。

通常设初始分辨率为 2^0，原始信号记为 $a_0(n)$，按分解算法逐次降低分辨率将信号分解为离散逼近信号和离散细节信号，如图 3.1 所示。设 a_0 有 N 个样本，将它分解为 a_1 和 d_1，由于滤波后仅取偶数样本，故 a_1 和 d_1 各为 $N/2$ 个样本，总样本数仍为 N，令 $1\leqslant j\leqslant J$，则分辨率为 2^{-j} 时，a_j 和 d_j 各有 $2^{-j}N$ 个样本。这样，如下一组离散信号

$$\{a_J;d_j,1\leqslant j\leqslant J\} \tag{3.2.23}$$

总共仍有 N 个样本。按 Mallat 算法进行信号分解后总样本数不变，这是由于采用了正交规范基的结果，意味着既没丢失信息，又不存在信息冗余。

3.2.2　重构算法

现在讨论如何从(3.2.22)式所示的一组离散信号重构 a_L。

由于 $V_j=V_{j+1}\oplus W_{j+1}$，而 V_{j+1} 和 W_{j+1} 的正交规范基分别为函数族 $\{\phi_{j+1,k}(t)\}_{k\in Z}$ 和 $\{\psi_{j+1,k}(t)\}_{k\in Z}$，故 $\{\phi_{j+1,k}(t),\psi_{j+1,k}(t)\}_{k\in Z}$ 是 V_j 的正交规范基。因为 $\phi_{j,n}(t)\in V_j$，所以可将它展开为

$$\phi_{j,n}(t)=\sum_k\langle\phi_{j,n}(u),\phi_{j+1,k}(u)\rangle\phi_{j+1,k}(t)+\sum_k\langle\phi_{j,n}(u),\psi_{j+1,k}(u)\rangle\psi_{j+1,k}(t)$$

上式中，由内积表示的第一个系数为

$$\langle\phi_{j,n}(u),\phi_{j+1,k}(u)\rangle=\int_{-\infty}^{+\infty}2^{-\frac{j}{2}}\phi(2^{-j}u-n)2^{-\frac{j+1}{2}}\phi(2^{-(j+1)}u-k)\mathrm{d}u$$

$$=2^{\frac{1}{2}}\int_{-\infty}^{+\infty}\phi(y-k)\phi(2y-n)\mathrm{d}y=\sqrt{2}\int_{-\infty}^{+\infty}\phi(x)\phi(2x+2k-n)\mathrm{d}x$$

由 $h(n)$ 的定义(3.2.5)式得

$$\langle\phi_{j,n}(u),\phi_{j+1,k}(u)\rangle=h(n-2k)$$

类似的，展开式的第二个系数为

$$\langle\phi_{j,n}(u),\psi_{j+1,k}(u)\rangle=g(n-2k)$$

于是展开式可写成

$$\phi_{j,n}(t)=\sum_k h(n-2k)\phi_{j+1,k}(t)+\sum_k g(n-2k)\psi_{j+1,k}(t)$$

将上式两边与$f(t)$求内积，并注意到(3.2.4)式和(3.2.16)式，则有

$$a_j(n) = \sum_k h(n-2k)a_{j+1}(k) + \sum_k g(n-2k)d_{j+1}(k) \tag{3.2.24}$$

这就是重构算法的数学表达式，如图3.2所示。

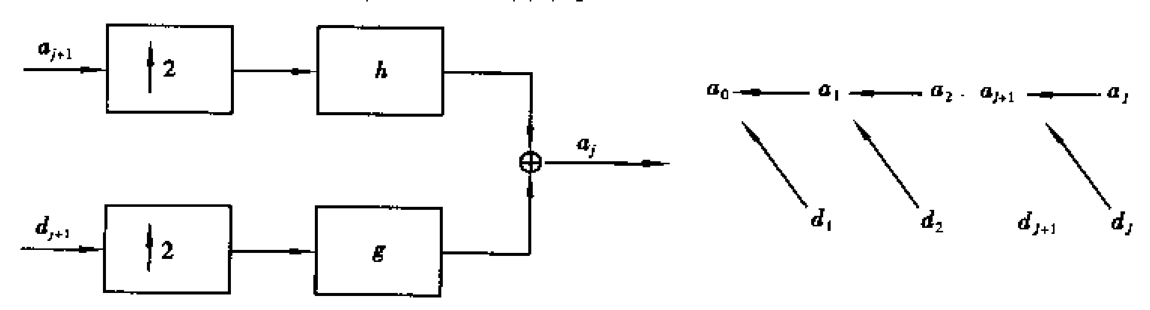


图3.2 Mallat信号重构算法

图3.2中，符号↑2表示上取样，即将a_{j+1}（或d_{j+1}）拉长，置奇数时刻信号值为零，偶数时刻保留原来的信号值，这与信号分解时的下取样相对应。

现在解释上取样。令a_{j+1}的上取样为a'_{j+1}，则a'_{j+1}可写成

$$a'_{j+1}(l) = \begin{cases} a_{j+1}(l/2) & l = 2k \\ 0 & l = 2k+1 \end{cases} \qquad k = 0, \pm 1, \cdots$$

这样，a'_{j+1}通过滤波器$h(n)$后的输出表示为如下卷积和

$$\sum_l h(n-l)a'_{j+1}(l) = \sum_k h(n-2k)a_{j+1}(k)$$

这正是(3.2.23)式右边第1项。同理可得第2项。

3.2.3 离散小波基

在多分辨率分析中，$\{\phi_{j,n}(t)\}_{n\in Z}$为$V_j$的正交规范基，即

$$\langle \phi_{j,n}(t), \phi_{j,k}(t) \rangle = \int_{-\infty}^{+\infty} \phi_{j,n}(t)\phi_{j,k}(t)\mathrm{d}t = \delta_{n,k}$$

令$j=0, n=k=0$，知尺度函数$\phi(t)$应满足

$$\int_{-\infty}^{+\infty} \phi^2(t)\mathrm{d}t = 1$$

即尺度函数是规范化的，这意味着当$t\to\pm\infty$时，$\phi(t)$趋近于零。实际应用中，要求$\phi(t)$在时域是局部化的，即当$t\to\pm\infty$时，$\phi(t)$应很快地衰减到零，于是其面积应当是有限的，一般要求

$$\int_{-\infty}^{+\infty} \phi(t)\mathrm{d}t = \pm 1$$

显然$\phi(t)$不是小波，因为小波的面积必须为零。

用$\hat{\phi}(\omega)$表示$\phi(t)$的傅立叶变换，由上式知$|\hat{\phi}(\omega)|_{\omega=0}=1$。实际应用中，还要求$\omega\to\pm\infty$时，$\hat{\phi}(\omega)$也很快地衰减到零，也就是说，$\hat{\phi}(\omega)$可以看成一个低通滤波器的频率响应。和小波函数一样，尺度函数在时-频域内也是同时局部化的，但尺度函数是一个低通滤波器，而小波函数是一个带通滤波器。不难想到，尺度函数空间中包含着信号的低频信息，而小波函数空间中包含着信号的高频信息。将尺度函数空间V_j分解为V_{j+1}和W_{j+1}，意味着将$f_{A_j}(t)$分解为逼近信号$f_{A_{j+1}}(t)$和细节信号$f_{D_{j+1}}(t)$，$f_{A_{j+1}}(t)$包含着$f_{A_j}(t)$的低频部分，而$f_{D_{j+1}}(t)$则包含着$f_{A_j}(t)$的高频部分。

如定义 $\phi_j(t)=2^{-\frac{j}{2}}\phi(2^{-j}t)$，$\overline{\phi}_j(t)=\phi_j(-t)$，由(3.2.4)式知离散逼近信号

$$\begin{aligned} a_j(n) &= \int_{-\infty}^{+\infty} f(u)\phi_{j,n}(u)\mathrm{d}t = 2^{-\frac{j}{2}}\int_{-\infty}^{+\infty} f(u)\phi(2^{-j}u-n)\mathrm{d}u \\ &= [f(t) * \overline{\phi}_j(t)]_{t=2^jn} \end{aligned}$$

也就是说，$a_j(n)$ 是 $f(t)$ 和 $\overline{\phi}_j(t)$ 卷积之后的取样，取样间隔为 2^j，即每单位长度的样本数为 2^{-j}。$\overline{\phi}_j(t)$ 的傅立叶变换为 $2^{j/2}\hat{\phi}^*(2^j\omega)$，仍是一个低通滤波器，所以 $a_j(n)$ 可理解为 $f(t)$ 通过低通滤波器之后取样的结果，也就是说，离散逼近信号描述了信号的低频成分。

类似的，定义 $\psi_j(t)=2^{-j/2}\psi(2^{-j}t)$，$\overline{\psi}_j(t)=\psi_j(-t)$，由(3.2.16)知离散细节信号

$$\begin{aligned} d_j(n) &= \int_{-\infty}^{+\infty} f(u)\psi_{j,n}(u)\mathrm{d}u = 2^{-\frac{j}{2}}\int_{-\infty}^{+\infty} f(u)\psi(2^{-j}u-n)\mathrm{d}u \\ &= [f(t) * \overline{\psi}_j(t)]_{t=2^jn} \end{aligned}$$

由于 $\overline{\psi}_j(t)$ 的傅立叶变换为 $2^{\frac{j}{2}}\hat{\psi}^*(2^j\omega)$，仍然是一个带通滤波器，故 $d_j(n)$ 是信号 $f(t)$ 通过带通滤波器之后取样的结果，也就是说，离散细节信号描述了信号的高频成分。综上所述，用 Mallat 算法对信号进行逐级分解时，逼近信号保留了原始信号的大致变化趋势，而细节信号则突出了原始信号的快速变化。

设原始离散信号为 a_L，它对应的尺度为 2^L，在尺度 $2^L\leqslant 2^j\leqslant 2^J$ 上用 Mallat 算法对它进行逐级分解时，得到如下一组离散信号

$$\{a_J; d_j, L\leqslant j\leqslant J\}$$

其中，离散逼近和离散细节信号可写作

$$a_j(n) = \langle a_L(k), \phi_j(k-2^{j-L}n)\rangle = a_L * \overline{\phi}_j(2^{j-L}n)$$

和

$$d_j(n) = \langle a_L(k), \psi_j(k-2^{j-L}n)\rangle = a_L * \overline{\psi}_j(2^{j-L}n)$$

这实际上是将信号 a_L 在矢量族

$$[\{\phi_J(k-2^{J-L}n)\}_{n\in Z}, \{\psi_j(k-2^{j-L}n)\}_{L<j\leqslant J, n\in Z}] \tag{3.2.25}$$

进行分解。对于共轭镜像滤波器，可以证明这个族是 $l^2(Z)$ 的一组正交规范基。这些离散的矢量近似于对连续时间尺度函数 $\phi_j(t)$ 和小波 $\psi_j(t)$ 的均匀采样。

例 3.1　分段常数逼近。

$[0,1)$ 区间的特征函数

$$I_{[0,1)}(t) = \begin{cases} 1, & 0\leqslant t<1 \\ 0, & \text{其他} \end{cases} \tag{3.2.26}$$

是生成多分辨率逼近的一个尺度函数，$\phi(t)=I_{[0,1)}(t)$，V_j 是在每段区间 $[2^jn, 2^j(n+1))_{n\in Z}$ 内为常数的分段函数空间，每段区间的长度均为 2^j。因为那些在长度为 2^{j+1} 的区间上为常数的函数在长度为 2^j 的区间上也是常数，所以有 $V_{j+1}\subset V_j$。其余的多分辨率逼近的性质的证明留给读者。

由于

$$\phi_{j,n}(t) = 2^{-\frac{j}{2}}\phi(2^{-j}t-n) = \begin{cases} 2^{-\frac{j}{2}}, & 2^jn\leqslant x<2^j(n+1) \\ 0, & \text{其他} \end{cases}$$

所以

$$a_j(n) = 2^{-\frac{j}{2}}\int_{2^jn}^{2^j(n+1)} f(u)\mathrm{d}u = 2^{\frac{j}{2}}f_{2^jn}$$

其中f_{2^jn}表示$f(t)$在区间$[2^jn,2^j(n+1))_{n\in Z}$上的平均值。这时逼近信号

$$A_j(t) = \sum_n f_{2^jn}I_{[2^jn,2^j(n+1))}(t)$$

是分段常数函数,在区间$[2^jn,2^j(n+1))_{n\in Z}$内的值为f_{2^jn}。j越小,每段区间的长度2^j越小,$A_j(t)$对$f(t)$的逼近程度越高。

这时小波为 Harr 函数

$$\psi(t) = \begin{cases}1,0 \leqslant t < 1/2\\ -1,1/2 \leqslant t < 1\\ 0,其他\end{cases} \tag{3.2.27}$$

例 3.2 Shannon 逼近。

分段常数逼近中的尺度函数和小波在时域是紧支的。与此对应,频率带限函数也产生多分辨率逼近。在 Shannon 多分辨率逼近中,尺度函数的傅立叶变换是区间$[-\pi,\pi]$上的特征函数,$\hat{\phi}(\omega)=I_{[-\pi,\pi]}(\omega)$,它是一个理想低通滤波器,尺度函数

$$\phi(t) = \frac{\sin\pi t}{\pi t} \tag{3.2.28}$$

因为

$$\langle\phi(t),\phi(t-n)\rangle = \int_{-\infty}^{+\infty}\phi(t)\phi(t-n)\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{\phi}(\omega)\hat{\phi}(\omega)\mathrm{e}^{i\omega n}\mathrm{d}\omega$$
$$= \frac{1}{2\pi}\int_{-\pi}^{+\pi}\mathrm{e}^{i\omega n}\mathrm{d}\omega = \begin{cases}1,n=0\\0,n\neq 0\end{cases}$$

所以$\{\phi(t-n)\}_{n\in Z}$构成正交规范基。对傅立叶变换的支集为$[-\pi,\pi]$的任意一个函数$g(t)$,有

$$\langle g(t),\phi(t-n)\rangle = \int_{-\infty}^{+\infty}g(t)\phi(t-n)\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{g}(\omega)\hat{\phi}(\omega)\mathrm{e}^{i\omega n}\mathrm{d}\omega$$
$$= \frac{1}{2\pi}\int_{-\pi}^{+\pi}\hat{g}(\omega)\mathrm{e}^{i\omega n}\mathrm{d}\omega = g(n)$$

其中,$g(n)$可解释为对$g(t)$采样的结果(采样周期为 1)。于是$g(t)$表示为

$$g(t) = \sum_n g(n)\phi(t-n) = \sum_n g(n)\frac{\sin\pi(t-n)}{\pi(t-n)} \tag{3.2.29}$$

这正是众所周知的采样定理。

定义V_j为如下函数的集合:它们是其傅立叶变换的支集为$[-2^{-j}\pi,2^{-j}\pi]$的频率带限函数。假定$g(t)$的傅立叶变换的支集为$[-2^{-(j+1)}\pi,2^{-(j+1)}\pi]$,即$g(t)\in V_{j+1}$,那么它的傅立叶变换的支集也必然属于$[-2^{-j}\pi,2^{-j}\pi]$,即$g(t)\in V_j$,所以$V_{j+1}\subset V_j$。这就证明了多分辨率逼近的函数空间的因果性。其余性质的证明也同样留给读者。

上面已证明$\{\phi(t-n)\}_{n\in Z}$是V_0的正交规范基。类似的,也可证明$\{\phi_{j,n}(t)\}_{n\in Z}$是$V_j$的正交规范基。且

$$a_j(n) = \langle f(t),\phi_{j,n}(t)\rangle = \int_{-\infty}^{+\infty}f(t)\phi_{j,n}(t)\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(\omega)2^{\frac{j}{2}}\hat{\phi}(2^j\omega)\mathrm{e}^{i2^j\omega n}\mathrm{d}\omega$$

$$= \frac{2^{\frac{j}{2}}}{2\pi}\int_{-2^{-j}\pi}^{2^{-j}\pi}\hat{f}(\omega)\mathrm{e}^{i2^j\omega n}\mathrm{d}\omega = 2^{\frac{j}{2}}f(2^jn)$$

所以

$$f_{A_j}(t) = A_jf(t) = \sum_n a_j(n)\phi_{j,n}(t) = \sum_n f(2^jn)\phi(2^{-j}t-n)$$
$$= \sum_n f(2^jn)\frac{\sin\pi(2^{-j}t-n)}{\pi(2^{-j}t-n)}$$

如果$f(t)\in V_j$，即其傅立叶变换$\hat{f}(\omega)$的支集为$[-2^{-j}\pi,2^{-j}\pi]$，则$A_jf(t)=f(t)$，上式就是采样定理的表达式，采样周期为2^j。对于具有无限带宽的信号$f(t)$，当$j\to-\infty$时，支集$[-2^{-j}\pi,2^{-j}\pi]$将覆盖整个频率轴，$A_jf(t)$将越来越逼近$f(t)$。

例 3.5 将证明，这时小波的傅立叶变换为

$$\hat{\psi}(\omega) = \begin{cases} \mathrm{e}^{-i\omega/2}, & \omega\in[-2\pi,-\pi]\cup[\pi,2\pi] \\ 0, & 其他 \end{cases}$$

它是一个理想带通滤波器，通带为$[-2\pi,-\pi]\cup[\pi,2\pi]$。由上式可得 Shannon 小波

$$\psi(t) = \frac{\sin 2\pi(t-1/2)-\sin\pi(t-1/2)}{\pi(t-1/2)} \tag{3.2.30}$$

Shannon 多分辨率逼近实际上是对频率轴的一种划分，如图 3.3 所示。假若$f(t)$是带限信号，其傅立叶变换的支集为$[-\pi,\pi]$，即$f(t)\in V_0$，意味着$a_0(n)=f(n)$完全描述了信号的特征。进行信号分解时，逼近信号$A_jf(t)$包含了$f(t)$在频带$[-2^{-j}\pi,2^{-j}\pi]$内的信息，而细节信号$D_jf(t)$则包含了$f(t)$在频带$[-2^{1-j}\pi,-2^{-j}\pi]\cup[2^{-j}\pi,2^{1-j}\pi]$内的信息。

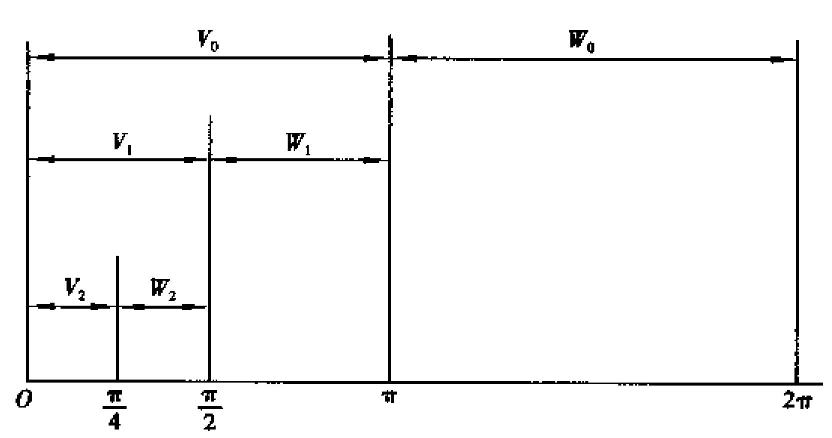


图 3.3　Shannon 逼近的频域划分

Shannon 逼近说明多分辨率分析实质上是对频率域的划分，这是具有普遍意义的。对其他多分辨率逼近，频域划分也基本上与 Shannon 多分辨率逼近类似，只不过逼近信号与细节信号的频带之间有部分重叠。图 3.4(a)和(b)为尺度函数和 Meyer 小波，傅立叶变换$\hat{\phi}(\omega)$的能量主要集中在$[-\pi,\pi]$上，如图 3.4(c)所示。像大多数正交小波一样，$\hat{\psi}(\omega)$的能量主要集中在$[-2\pi,-\pi]\cup[\pi,2\pi]$上，如图 3.4(d)所示。

对于生成$L^2(R)$的一组正交基的任何$\psi(t)$，可证

$$\sum_{j=-\infty}^{+\infty}|\hat{\psi}(2^j\omega)|^2 = 1, \forall\omega\in R-\{0\} \tag{3.2.31}$$

对一个确定的分辨率，$\hat{\psi}(2^j\omega)$是一个带通滤波器，$D_jf(t)$刻画了信号在频带$[-2^{1-j}\pi,-2^{-j}\pi]$

$\cup[2^{-j}\pi, 2^{1-j}\pi]$内的细节。由(3.2.31)式,有

$$\sum_{j=-\infty}^{+\infty} |\hat{f}(\omega)\hat{\psi}(2^j\omega)|^2 = |\hat{f}(\omega)|^2$$

这意味着可以从所有不同频带内的细节获得信号的全部信息。

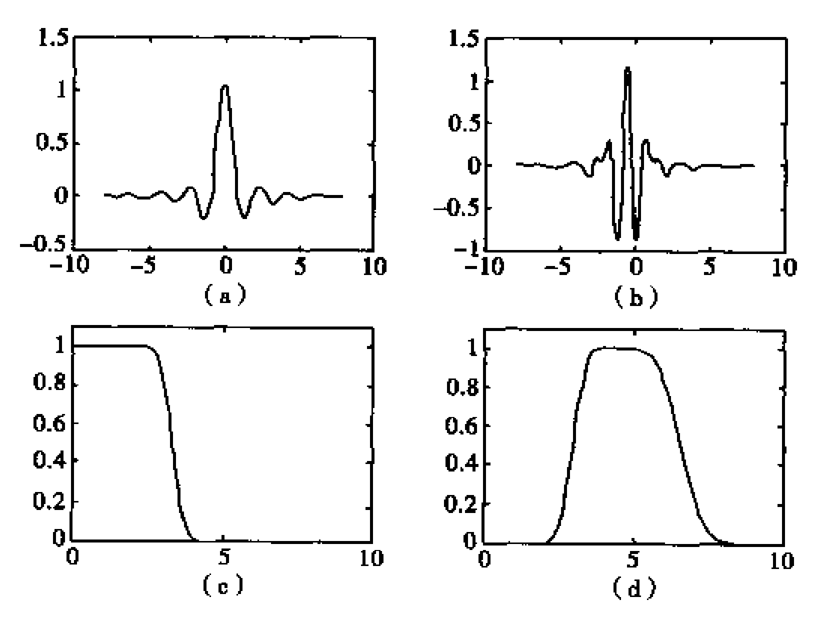


图 3.4 Meyer 小波的 (a) ϕ, (b) ψ, (c) $\hat{\phi}$, (d) $\hat{\psi}$

例 3.3 样条逼近。

分段常数逼近不光滑,用来逼近光滑函数效率很差。样条逼近是用分段多项式来逼近光滑函数。$m \geqslant 0$ 阶样条空间 V_j 是如下函数的集合:在区间$[2^jn, 2^j(n+1)]_{n\in Z}$上是 m 次多项式,$m-1$阶连续可微。$m=0$ 时,就是分段常数逼近。$m=1$ 时,V_j 中的函数是分段线性的,连续的。

多项式样条的 Riesz 基可通过基数 B 样条构造。m 次 B 样条 $\theta_m(t)$的傅立叶变换是

$$\hat{\theta}_m(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2}\right)^{m+1} e^{-\frac{i\varepsilon\omega}{2}} \tag{3.2.32}$$

如果 m 是奇数,$\varepsilon=0$,$\theta_m(t)$的支集为$[-(m+1)/2,(m+1)/2]$,对称中心为 0。如果 m 是偶数,$\varepsilon=1$,$\theta_m(t)$的支集为$[-m/2,(m/2)+1]$,对称中心为 1/2。可以证明$\{\theta_m(t-n)\}_{n\in Z}$是 V_0 的 Riesz 基。通过正交化可以得到正交规范基,为此将上式代入(3.2.1)式得

$$\hat{\phi}(\omega) = \frac{\exp(-i\varepsilon\omega/2)}{\omega^{m+1}\sqrt{S_{2m+2}(\omega)}} \tag{3.2.33}$$

其中

$$S_l(\omega) = \sum_{k=-\infty}^{+\infty} \frac{1}{(\omega+2k\pi)^l} \tag{3.2.34}$$

$S_{2m+2}(\omega)$的封闭表达式可通过计算下面恒等式的 $2m$ 阶导数得到

$$S_2(\omega) = \frac{1}{4\sin^2(\omega/2)}$$

对线性样条,$m=1$ 且

$$S_4(\omega) = \frac{1+2\cos^2(\omega/2)}{48\sin^4(\omega/2)}$$

从而,有

$$\hat{\phi}(\omega) = \frac{4\sqrt{3}\sin^2(\omega/2)}{\omega^2\sqrt{1+2\cos^2(\omega/2)}}$$

三次样条尺度函数对应于 $m=3$,将

$$S_8(\omega) = \frac{5+30\cos^2(\omega/2)+30\sin^2(\omega/2)\cos^2(\omega/2)}{105\ 2^8\sin^8(\omega/2)} + \frac{70\cos^4(\omega/2)+2\sin^4(\omega/2)\cos^2(\omega/2)+2/3\sin^6(\omega/2)}{105\ 2^8\sin^8(\omega/2)}$$

代入(3.2.33)式,可算出 $\hat{\phi}(\omega)$。

3.3 尺度滤波器

从上述分析可以看到,Mallat 算法中出现的是离散逼近信号和离散细节信号,以及数字滤波器 $h(n)$ 和 $g(n)$,而没有出现模拟信号、尺度函数和小波。其实多分辨率分析应当从连续和离散两个方面来理解,它既是函数空间和模拟信号在相邻分辨率之间的逐级分解,又是离散逼近信号在相邻分辨率之间的逐级分解。而且模拟和离散这两方面是彼此联系的。

在 Mallat 算法中,出现了 4 个数字滤波器 $h(n)$、$g(n)$ 和 $\bar{h}(n)$、$\bar{g}(n)$,即一个滤波器组。但是 $\bar{h}(n)$、$\bar{g}(n)$ 是 $h(n)$、$g(n)$ 的镜像,而 $g(n)$ 可由 $h(n)$ 导出,故 Mallat 算法中,起核心作用的是 $h(n)$。

多分辨率逼近的因果性质(3.1.3)式要求 $V_0 \subset V_{-1}$,故 $\phi(t)$ 可按 V_{-1} 的正交规范基展开成

$$\phi(t) = \sqrt{2}\sum_n h(n)\phi(2t-n) \tag{3.3.1}$$

其中

$$h(n) = \sqrt{2}\int_{-\infty}^{+\infty}\phi(t)\phi(2t-n)\mathrm{d}t$$

(3.3.1)式称为二尺度方程,它说明尺度函数可以表示为它自身在更精细一级尺度上平移的线性组合,线性组合的系数就是尺度滤波器 $h(n)$。由尺度滤波器可以生成多分辨率逼近,而且也可以通过它构造正交小波。

对二尺度方程(3.3.1)式两边做傅立叶变换,有

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}}H\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \tag{3.3.2}$$

逐次使用上式进行迭代运算得

$$\hat{\phi}(\omega) = \left(\prod_{p=1}^{P}\frac{H(2^{-p}\omega)}{\sqrt{2}}\right)\hat{\phi}(2^{-P}\omega)$$

如果 $\hat{\phi}(\omega)$ 在 $\omega=0$ 连续,那么 $\lim\limits_{P\to+\infty}\hat{\phi}(2^{-P}\omega)=\hat{\phi}(0)$,所以

$$\hat{\phi}(\omega) = \left(\prod_{p=1}^{+\infty}\frac{H(2^{-p}\omega)}{\sqrt{2}}\right)\hat{\phi}(0) \tag{3.3.3}$$

下述定理给出上列无穷乘积是某个尺度函数的傅立叶变换的必要和充分条件。

定理 3.3 设 $\phi(t)\in L^2(R)$ 是一个可积的尺度函数,那么

$$h(n) = \sqrt{2}\int_{-\infty}^{+\infty} \phi(t)\phi(2t-n)\mathrm{d}t \tag{3.3.4}$$

它的傅立叶变换

$$H(\omega) = \sum_{n=-\infty}^{\infty} h(n)\mathrm{e}^{-i\omega n} \tag{3.3.5}$$

满足

$$|H(\omega)|^2 + |H(\omega+\pi)|^2 = 2, \forall\omega \in R \tag{3.3.6}$$

和

$$H(0) = \sqrt{2} \tag{3.3.7}$$

反之,如 $H(\omega)$ 在 $\omega=0$ 的某邻域内连续可微,满足(3.3.6)式和(3.3.7)式,且

$$\inf_{\omega\in[-\pi/2,\pi/2]} |H(\omega)| > 0 \tag{3.3.8}$$

则

$$\hat{\phi}(\omega) = \prod_{p=1}^{+\infty} \frac{H(2^{-p}\omega)}{\sqrt{2}} \tag{3.3.9}$$

是某个尺度函数 $\phi(t) \in L^2(R)$ 的傅立叶变换。

满足(3.3.6)式的数字滤波器称为共轭滤波器。由(3.3.6)式和(3.3.7)式可得 $|H(\pi)|=0$,所以 $h(n)$ 是一个低通滤波器。

由(3.2.10)式,$|G(0)|=|H(\pi)|=0$,$|G(\pi)|=|H(0)|=\sqrt{2}$,故 $g(n)$ 是一个高通滤波器。可以证明,$\{\psi_{j,n}(t)\}_{n\in Z}$ 是 W_j 的正交规范基的充分必要条件是

$$|G(\omega)|^2 + |G(\omega+\pi)|^2 = 2 \tag{3.3.10}$$

和

$$G(\omega)H^*(\omega) + G(\omega+\pi)H^*(\omega+\pi) = 0 \tag{3.3.11}$$

类似的,由于 $W_0 \subset V_{-1}$,故 $\psi(t)$ 可按 V_{-1} 的正交规范基展开成

$$\psi(t) = \sqrt{2}\sum_n g(n)\phi(2t-n) \tag{3.3.12}$$

其中

$$g(n) = \sqrt{2}\int_{-\infty}^{+\infty} \psi(t)\phi(2t-n)\mathrm{d}t \tag{3.3.13}$$

(3.3.12)式也是一个二尺度方程,它将小波表示为在更精细一级尺度上尺度函数平移的线性组合,线性组合的系数就是 $g(n)$。将(3.3.12)式的两边作傅立叶变换可得

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}G\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \tag{3.3.14}$$

例 3.4 对分段常数逼近,由(3.3.4)式可以算出 $h(0)=h(1)=1/\sqrt{2}$,其余的 $h(n)=0$。从而由(3.2.14)式可以算出 $g(n)$ 不等于 0 的两个系数分别为 $g(0)=1/\sqrt{2}$,$g(1)=-1/\sqrt{2}$。由(3.3.12)式得

$$\psi(t) = \phi(2t) - \phi(2t-1) = I_{[0,1/2]}(t) - I_{[1/2,1]}(t)$$

这正是(3.2.27)式所示的哈尔小波。由(3.3.5)式可以算出 $H(\omega)=\sqrt{2}\mathrm{e}^{-i\omega/2}\cos(\omega/2)$,同样可算出 $G(\omega)=\sqrt{2}\mathrm{e}^{i(\pi-\omega)/2}\sin(\omega/2)$。

例 3.5 对 Shannon 多分辨率逼近,$\hat{\phi}(\omega)=I_{[-\pi,\pi]}(\omega)$,它是一个理想低通滤波器,通带

为[$-\pi,\pi$]。由(3.3.2)式可得

$$H(\omega)=\sqrt{2}I_{[-\pi/2,\pi/2]}(\omega),\ \forall\omega\in[-\pi,\pi]$$

它是一个理想低通数字滤波器。由(3.2.10)式,有

$$\forall\omega\in[-\pi,\pi],G(\omega)=\begin{cases}\sqrt{2}\mathrm{e}^{-\mathrm{i}\omega}, & \omega\in[-\pi,-\pi/2]\cup[\pi/2,\pi]\\0, & \text{其他}\end{cases}$$

$$\hat{\psi}(\omega)=\begin{cases}\mathrm{e}^{-\mathrm{i}\omega/2}, \omega\in[-2\pi,-\pi]\cup[\pi,2\pi]\\0, & \text{其他}\end{cases}$$

可见 $G(\omega)$ 是一个理想高通数字滤波器,而 $\hat{\psi}(\omega)$ 是一个理想带通滤波器,通带为[$-2\pi,-\pi$] ∪ [$\pi,2\pi$]。对 $\hat{\psi}(\omega)$ 做傅立叶反变换,有

$$\begin{aligned}\psi(t)&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{\psi}(\omega)\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}\omega\\&=\frac{1}{2\pi}\int_{-2\pi}^{-\pi}\mathrm{e}^{\mathrm{i}\omega(t-1/2)}\mathrm{d}\omega+\frac{1}{2\pi}\int_{\pi}^{2\pi}\mathrm{e}^{\mathrm{i}\omega(t-1/2)}\mathrm{d}\omega\\&=\frac{\sin 2\pi(t-1/2)-\sin\pi(t-1/2)}{\pi(t-1/2)}\end{aligned}$$

这正是(3.2.30)式所给出的结果。

例 3.6　对样条多分辨率逼近,由(3.3.2)式可得尺度滤波器 $h(n)$ 的傅立叶变换

$$H(\omega)=\sqrt{2}\frac{\hat{\phi}(2\omega)}{\hat{\phi}(\omega)}$$

将(3.2.33)式代入便得到 m 次多项式样条对应的尺度滤波器的傅立叶变换

$$H(\omega)=\sqrt{\frac{S_{2m+2}(\omega)}{2^{2m+1}S_{2m+2}(2\omega)}}\exp\left(\frac{-\mathrm{i}\varepsilon\omega}{2}\right)\tag{3.3.15}$$

对于线性样条, $m=1$,有

$$H(\omega)=\sqrt{2}\left[\frac{1+2\cos^2(\omega/2)}{1+2\cos^2\omega}\right]^{\frac{1}{2}}\cos^2(\omega/2)$$

表 3.1　3 次样条的尺度滤波器 $h(n)$

n	$h(n)$	n	$h(n)$	n	$h(n)$
0	0.766 130 398	7, −7	−0.017 982 291	14, −14	−0.001 103 748
1, −1	0.433 923 147	8, −8	0.008 685 294	15, −15	−0.000 927 187
2, −2	−0.050 201 753	9, −9	0.008 201 477	16, −16	0.000 559 952
3, −3	−0.110 036 987	10, −10	−0.004 353 840	17, −17	0.000 462 093
4, −4	0.032 080 869	11, −11	−0.003 882 426	18, −18	−0.000 285 414
5, −5	0.042 068 328	12, −12	0.002 186 714	19, −19	−0.000 232 304
6, −6	−0.017 176 331	13, −13	0.001 882 120	20, −20	0.000 146 098

将(3.2.32)式和(3.3.15)式代入(3.3.14)式可得

$$\hat{\psi}(\omega) = \frac{e^{-i\frac{\omega}{2}}e^{i\frac{\epsilon\pi}{2}}}{\omega^{m+1}}\sqrt{\frac{S_{2m+2}(\omega/2+\pi)}{S_{2m+2}(\omega)S_{2m+2}(\omega/2)}} \tag{3.3.16}$$

它被称为 Battle-Lemarié 小波。对 m 次样条，$H(\omega)$ 及其前 m 阶导数在 $\omega=\pi$ 处等于零，由定理 4.1 知 $\psi(t)$ 有 $m+1$ 阶消失矩。由于 $\psi(t)$ 是 m 次多项式样条，所以它 $m-1$ 次连续可微。当 m 为奇数时，$\psi(t)$ 关于 1/2 对称；当 m 为偶数时，$\psi(t)$ 关于 1/2 反对称。这时 $\phi(t)$ 和 $\psi(t)$ 都不是紧支撑的，相应的 $h(n)$ 和 $g(n)$ 是无限冲激响应数字滤波器。表 3.1 列出了 $m=3$ 时尺度滤波器的系数。

从上述分析可以看到，在 Mallat 算法中，信号分解用到两个数字滤波器 $\bar{h}(n)$ 和 $\bar{g}(n)$，其中 $\bar{h}(n)$ 是低通滤波器，而 $\bar{g}(n)$ 是高通滤波器；信号重构用到两个数字滤波器 $h(n)$ 和 $g(n)$，其中 $h(n)$ 是低通滤波器，而 $g(n)$ 是高通滤波器。

在 Matlab 中，函数 Orthfilt 用来计算这 4 个滤波器。其中如下二尺度(twin-scale)方程定义了所谓的尺度滤波器 W_n：

$$\frac{1}{2}\phi\left(\frac{x}{2}\right) = \sum_{n\in Z} W_n\phi(x-n)$$

不难看到，除归一化系数外，上式与(3.3.1)式是一致的。若 $\phi(x)$ 是紧支的，那么 W_n 是一个长度为 $2N$ 的 FIR 低通数字滤波器，其和为 1，即

$$\sum_n W_n = 1$$

其等价关系为 $$W(\omega)|_{\omega=0} = 1$$

而且，其范数为 $1/\sqrt{2}$。然后经下列计算得到 4 个滤波器

$$\begin{array}{c} W \\ \downarrow \\ L_o_R = \dfrac{W}{\text{norm}(W)} \rightarrow L_o_D = \text{wrev}(L_o_R) \\ \downarrow \\ H_i_R = \text{qmf}(L_o_R) \rightarrow H_i_D = \text{wrev}(H_i_R) \end{array}$$

其中 L_o_R 就是 Mallat 算法中的 $h(n)$，而 L_o_D 就是 Mallat 算法中的 $\bar{h}(n)$。只不过与(3.2.7)式比较，为使 $\bar{h}(n)$ 保持因果性，除镜像对称外，再加上延迟，即

$$\bar{h}(n) = h[-(n-(2N-1))] = h(2N-1-n)$$

就是将 $h(n)$ 倒置得到 $\bar{h}(n)$。H_i_R 为 L_o_R 的正交镜像滤波器(Quadrature Mirror Filter)

$$H_i_R(n) = (-1)^n L_o_R(2N-1-n)$$

上式与(3.2.14)式是一致的，只是将(3.2.14)式取负号，而且延迟 $(2N-2)$ 便得到上式

$$\begin{aligned} g(n) &= -(-1)^{1-n}h(1-n) = (-1)^n h(1-n) \xrightarrow{2N-2} \\ &\rightarrow (-1)^{n-(2N-2)}h[1-(n-(2N-2))] = (-1)^n h(2N-1-n) \end{aligned}$$

同样的，延迟也是为了使 $g(n)$ 是因果性的。

函数 Orthfilt 返回分解低通(L_o_D)、分解高通(H_i_D)及重构低通(L_o_R)、重构高通(H_i_R)。图 3.5(b)、(c)、(d)、(e)分别画出了与 db2 相联系的这 4 个滤波器，图 3.5(a)为尺度滤波器 W_n。图 3.5(f)、(g)则分别画出了低通滤波器和高通滤波器的频率特性，其中横轴为归一化数字频率，1.0 对应数字频率 π 和奈奎斯特频率(Hz)。

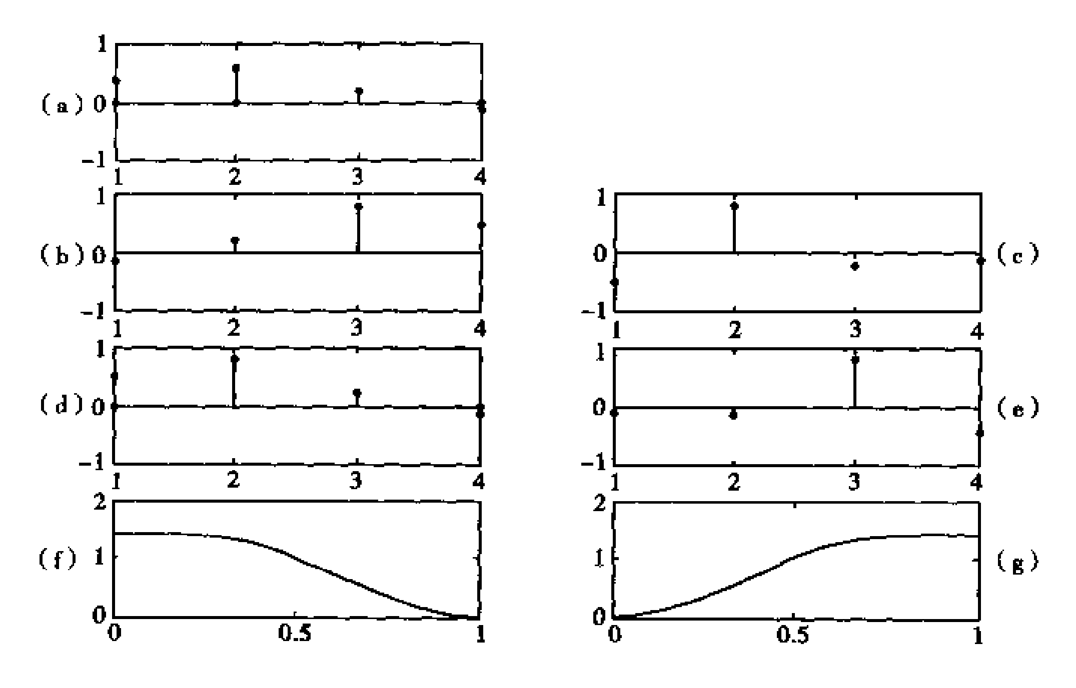


图3.5　db2滤波器组

3.4　Mallat算法的实现

3.4.1　信号初始化

Mallat算法实际上是对离散信号逐级迭代计算的过程。那么原始信号 $a_L(n)$ 是如何得到的呢？如果待处理的信号本身就是数字信号，则直接把它作为 $a_L(n)$。如果待处理的信号是模拟信号，则按采样定理对它采样，把采样得到的离散样本作为 $2^{-L/2}a_L(n)$。按离散逼近信号的定义，有

$$2^{-\frac{L}{2}}a_L(n) = 2^{-\frac{L}{2}}\langle f(t),\phi_{L,n}(t)\rangle = \int_{-\infty}^{+\infty} f(t)\ \frac{1}{2^L}\phi\left(\frac{t-2^L n}{2^L}\right)\mathrm{d}t$$

它可以解释为信号在一个与尺度 2^L 成比例的区域内的平均。实际上，大多数离散输入信号都是用有限分辨率的器件对模拟信号平均和采样得到的。例如CCD摄像机中的每一光敏单元都在一定范围内平均输入光强。在上列积分中，对足够精细的尺度 $T_s=2^L$，尺度函数的持续时间很短，但其面积为1，故可看成一个位于 $t=2^L n=nT_s$ 的脉冲函数，于是有

$$2^{-\frac{L}{2}}a_L(n) \approx f(nT_s) \tag{3.4.1}$$

也就是说，$2^{-L/2}a_L(n)$ 近似为对原始模拟信号的采样，采样频率 $1/T_s=2^{-L}$。

类似小波函数的消失矩，也可以定义尺度函数的消失矩。如尺度函数满足

$$\int_{-\infty}^{+\infty}\phi(t)\mathrm{d}t = 1 \text{ 以及 } \int_{-\infty}^{+\infty}t^k\phi(t)\mathrm{d}t = 0, 1 \leqslant k < K \tag{3.4.2}$$

则称尺度函数具有 K 阶消失矩。如果信号 $f(t)$ 在 $t=2^L n=nT_s$ 的邻域内是 C^k 光滑的，且 $k<K$，那么，$f(t)$ 直到 k 阶的泰勒展开式表明

$$2^{-\frac{L}{2}}a_L(n) \approx f(nT_s) + O(2^{(k+1)L}) \tag{3.4.3}$$

这意味着原始模拟信号的采样值$f(nT_s)$确实是$2^{-L/2}a_L(n)$的高阶近似。

按采样定理

$$f(t) = \sum_k f(kT_s)\operatorname{sinc}(t-kT_s),\operatorname{sinc}(t) = \frac{\sin\pi t/T_s}{\pi t/T_s} \tag{3.4.4}$$

所以，严格说来

$$a_L(n) = \langle f(t),\phi_{L,n}(t)\rangle = \sum_k f(kT_s)\int_{-\infty}^{+\infty}\operatorname{sinc}(t-kT_s)\phi_{L,n}(t)\,dt \tag{3.4.5}$$

但按此式计算$a_L(n)$很繁琐。如仍将$2^{-L/2}\phi_{L,n}(t)$近似为一个位于$t=2^Ln=nT_s$的脉冲函数，则上式简化为

$$\begin{aligned} a_L(n) &= 2^{L/2}\sum_k f(kT_s)\operatorname{sinc}(2^Ln-kT_s) \\ &= 2^{L/2}\sum_k f(kT_s)\operatorname{sinc}((n-k)T_s) \end{aligned} \tag{3.4.6}$$

由例3.2的讨论可以看出，这意味着用$A_jf(t)$来逼近$f(t)$[3]。

此外，尽管在不同的分辨率级次j，时间平移参数n都取整数，但它们对应的平移步长2^{L+j}是不一样的。例如$j=1$时的平移步长是原始信号的2倍，故当$a_{L+1}(n)$和$d_{L+1}(n)$中的时间平移参数n取某一值n_0时，如以原始信号的时间尺度作参考，则时间平移量为$2n_0T_s$。

3.4.2 边界延拓

如果原始信号为有限长的时间序列，作卷积运算时在信号的两端会出现边界问题。

假定$a_L(n)$的持续时间为$0\leqslant n\leqslant N-1$，而分辨率为$j$级次时的样本数$N_j=2^{-j}N$，故$a_{L+j}(n)$和$d_{L+j}(n)$的持续时间为$0\leqslant n\leqslant N_j-1$。可用以下几种方法进行边界延拓。

(1)补零

就是置边界外的信号值等于0。这种方法最简单，但有可能在边界两端形成信号跳变，相当于在边界两端人为地引入了阶跃，从而使边界附近的小波系数幅值很大，但这并不是信号本身的特征。

(2)周期延拓

将有限长序列周期重复。例如对信号$a_{L+j}(n)$周期重复得到如下周期信号

$$\tilde{a}_{L+j}(n) = \sum_l a_{L+j}(n+lN_j) \tag{3.4.7}$$

周期延拓也有可能在边界两端形成信号跳变，从而使边界附近的小波系数幅值很大。

(3)对称延拓

$$a_{L+j}(n) = \begin{cases} a_{L+j}(-n), & -(N_j-1)<n<0 \\ a_{L+j}(2N_j-n-1), & N_j-1<n<2N_j-2 \end{cases} \tag{3.4.8}$$

在Matlab的小波工具箱中，除上述3种方式外，还有0阶和1阶插值方式。

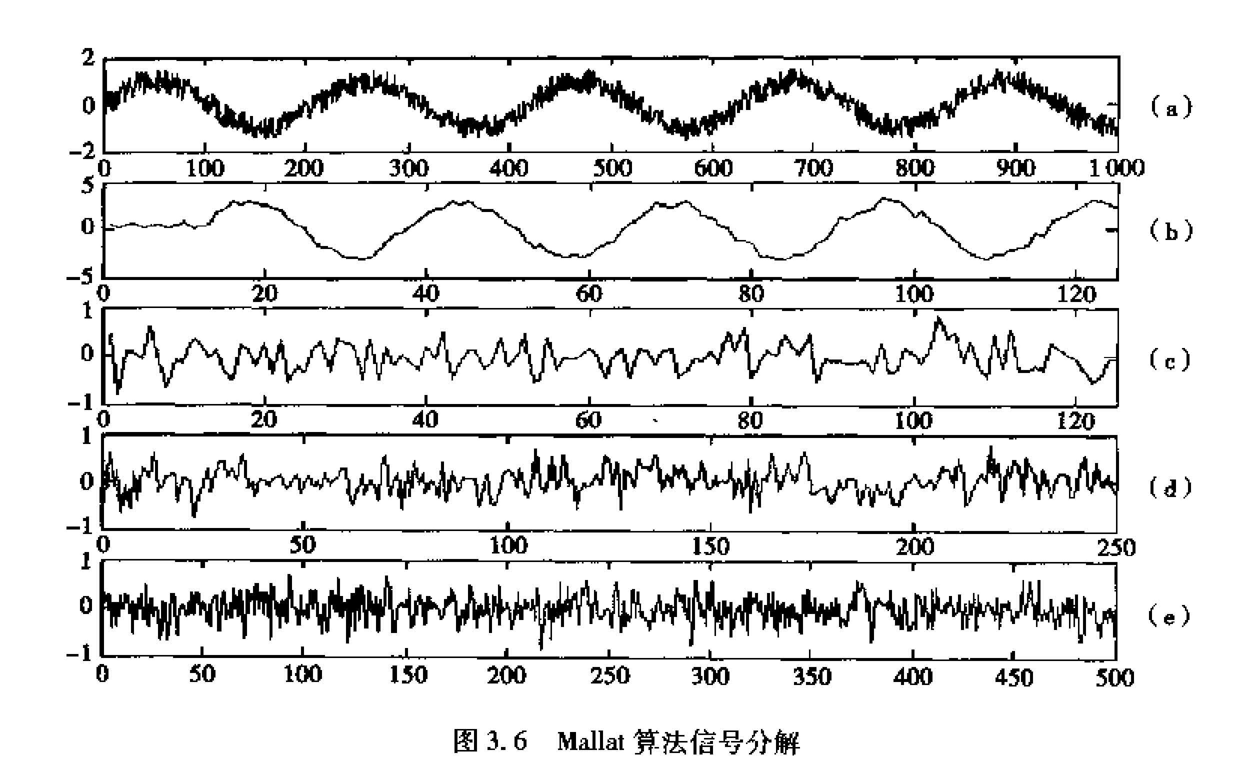


图3.6　Mallat算法信号分解

3.4.3　Mallat算法的实现

信号分解可以通过循环嵌套程序来实现，最外层为分辨率级次，$1 \leqslant j \leqslant J$；中间为时间变量，$0 \leqslant n \leqslant N_j - 1$；最里层为卷积求和变量 k。如果将信号分解的数学表达式(3.2.8)和(3.2.21)式改写为

$$a_{j+1}(n) = \sum_k \bar{h}(k) a_j(2n-k) \tag{3.4.9}$$

$$d_{j+1}(n) = \sum_k \bar{g}(k) a_j(2n-k) \tag{3.4.10}$$

或

$$a_{j+1}(n) = \sum_k h(k) a_j(2n+k) \tag{3.4.11}$$

$$d_{j+1}(n) = \sum_k g(k) a_j(2n+k) \tag{3.4.12}$$

当数字滤波器为有限冲击响应滤波器，且采用直接型结构实现时，上述4式的求和范围就很容易确定了。

在Matlab的小波工具箱中，函数wavedec可用来实现一维信号的多尺度分解。图3.6(a)为一含有白噪声的正弦信号，样本数为1 000。对它作3级分解，图3.6(b)为第3级离散逼近信号，它是原始信号3次低通滤波的结果，样本数为125。图3.6(c)、(d)、(e)依次为第3级、第2级、第1级的离散细节信号，样本数依次为125、250、500。我们已将各级离散细节信号和最后一级离散逼近信号的时间轴拉长，以便于以原始信号的时间尺度作为参考。如原始离散信号的单位间隔对应采样周期，则各级离散细节信号和最后一级离散逼近信号的时间平移参数，都可以对应模拟信号的时间。例如采样周期为1秒，原始离散逼近信号 $a_0(n)$ 中 $n=10$ 对应的时间为10秒。而在 $d_1(n)$ 中，$n=10$ 对应的时间为20秒。由图3.6可以看到，原始信号的低频成分(正弦分量)保留在离散逼近信号中，而其高频成分(噪声分量)则包含在离散细节

信号中。众所周知，白噪声具有很宽的频谱范围，所以3级离散细节信号的幅度基本一样。如果只保留第3级离散逼近信号，则既可去噪，又可达到数据压缩的目的。

3.4.4 区间上的小波基

区间上的小波基实际上是一个与边界延拓紧密相关的问题。定理3.2证明通过二进伸缩和平移可以构造出$L^2(R)$的小波基。对于有限区间上的信号，边界延拓实际上是扩展了信号的定义域，仍然用$L^2(R)$上的小波基来进行分解。也可以通过修改$L^2[R]$的小波基，来构造有限区间上的小波基，从而实现定义在有限区间上的信号的小波分解。下面讨论定义在区间[0,1]上的信号的分解，即构造$L^2[0,1]$的小波基的问题。对任意区间$[a,b]$，$L^2[a,b]$上的小波基可以对$L^2[0,1]$的小波基用b-a进行伸缩，用a进行平移得到。

构造区间上的小波基和边界延拓一样，也是为了避免在边界附近产生大幅值的小波系数。从$L^2(R)$的小波基$\{\psi_{j,n}\}_{j,n\in Z}$构造区间上的小波基时，如小波$\psi_{j,n}$的支集包含在区间[0,1]内，称为内小波，不经修改直接作为区间上的小波基；如小波$\psi_{j,n}$的支集覆盖了边界点$t=0$或$t=1$，称为边界小波，则要修改为支集包含在区间[0,1]内的小波。对尺度函数也要做类似的修改，这样在一个小于1的尺度$2^J<1$上的2^{-J}个尺度函数和在每个尺度$2^j\leqslant 2^J$上的2^{-j}个小波就构成了区间[0,1]的小波基

$$[\{\phi_{J,n}^{\text{int}}\}_{0\leqslant n<2^{-J}},\{\psi_{j,n}^{\text{int}}\}_{-\infty<j\leqslant J,0\leqslant n<2^{-j}}] \tag{3.4.13}$$

对定义在区间[0,1]上的连续时间信号采样，得到有N个离散时间样本的原始信号$a_L(n)$，对应的尺度为2^L，在尺度$2^L<2^j\leqslant 2^J\leqslant 1$上对$a_L(n)$进行分解，和$l^2(Z)$的情况类似，离散逼近和离散细节信号也可以写成离散内积

$$a_j(n)=\langle a_L(k),\phi_{j,n}^{\text{int}}(k)\rangle \text{ 和 } d_j(n)=\langle a_L(k),\psi_{j,n}^{\text{int}}(k)\rangle \tag{3.4.14}$$

同样可以证明

$$[\{\phi_{J,n}^{\text{int}}(k)\}_{0\leqslant n<2^{-J}},\{\psi_{j,n}^{\text{int}}(k)\}_{L<j\leqslant J,0\leqslant n<2^{-j}}]$$

构成C^N的正交规范基。

3.5 二维多分辨率分析

将一维多分辨率分析直接推广到二维情况，以便将多分辨率分析用于图像处理。与定义3.1类似，对$L^2(R^2)$的多分辨逼近是构造嵌套的闭子空间序列$\{V_j\}_{j\in Z}$，满足同样的因果性、完备性和二尺度关系等性质。图像$f(x,y)$在分辨率2^{-j}时的逼近定义为$f(x,y)$在$L^2(R^2)$的子空间$\tilde{V}_j$上的正交投影。

我们将讨论一种特定的情况，这时二维矢量空间$\tilde{V}_j$可以分离为两个相同的一维矢量空间V_j的张量积，即

$$\tilde{V}_j=V_j\otimes V_j,\forall j\in Z \tag{3.5.1}$$

如$\{V_j\}_{j\in Z}$为$L^2(R)$的多分辨率逼近，那么$\{\tilde{V}_j\}_{j\in Z}$构成对$L^2(R^2)$的多分辨率逼近。这时二维尺度函数为两个相同的一维尺度函数的乘积，即

$$\Phi(x,y)=\phi(x)\phi(y) \tag{3.5.2}$$

$\tilde{V}_j$ 的正交基为

$$\begin{aligned}&\{2^{-j}\Phi(2^{-j}x-n,2^{-j}y-m)\}_{n,m\in Z}\\ &=\{2^{-j/2}\phi(2^{-j}x-n)2^{-j/2}\phi(2^{-j}y-m)\}_{n,m\in Z}\\ &=\{\phi_{j,n}(x)\phi_{j,m}(y)\}_{n,m\in Z}\end{aligned} \tag{3.5.3}$$

上式中 j 为尺度参数,n 和 m 分别为沿 x 和 y 方向的平移参数。离散逼近信号定义为下述内积

$$a_j(n,m)=\langle f(x,y),\phi_{j,n}(x)\phi_{j,m}(y)\rangle \tag{3.5.4}$$

这时 $\tilde{V}_j$ 在 $\tilde{V}_{j-1}$ 中的正交补 $\tilde{W}_j$ 要稍微复杂一点。

定理 3.4　设 $\psi(x)$ 是与 $\phi(x)$ 相对应的一维小波,那么下列 3 个小波

$$\psi^1(x,y)=\phi(x)\psi(y),\psi^2(x,y)=\psi(x)\phi(y),\psi^3(x,y)=\psi(x)\psi(y) \tag{3.5.5}$$

经二进伸缩和平移形成的函数族

$$\begin{aligned}&\{2^{-j}\psi^1(2^{-j}x-n,2^{-j}y-m),2^{-j}\psi^2(2^{-j}x-n,2^{-j}y-m),\\ &2^{-j}\psi^3(2^{-j}x-n,2^{-j}y-m)\}_{n,m\in Z}\end{aligned} \tag{3.5.6}$$

是 $\tilde{W}_j$ 的正交规范基,且

$$\begin{aligned}&\{2^{-j}\psi^1(2^{-j}x-n,2^{-j}y-m),2^{-j}\psi^2(2^{-j}x-n,2^{-j}y-m),\\ &2^{-j}\psi^3(2^{-j}x-n,2^{-j}y-m)\}_{j,n,m\in Z}\end{aligned} \tag{3.5.7}$$

是 $L^2(R^2)$ 的正交规范基。

证明　$\tilde{W}_j$ 是 $\tilde{V}_j$ 在 $\tilde{V}_{j-1}$ 中的正交补,即

$$\tilde{V}_{j-1}=\tilde{V}_j\oplus\tilde{W}_j \tag{3.5.8}$$

由于 $V_{j-1}=V_j\oplus W_j$,将它代入(3.5.1)有

$$\begin{aligned}\tilde{V}_{j-1}&=V_{j-1}\otimes V_{j-1}=(V_j\oplus W_j)\otimes(V_j\oplus W_j)\\ &=(V_j\otimes V_j)\oplus(V_j\otimes W_j)\oplus(W_j\otimes V_j)\oplus(W_j\otimes W_j)\\ &=\tilde{V}_j\oplus\tilde{W}_j^1\oplus\tilde{W}_j^2\oplus\tilde{W}_j^3\end{aligned} \tag{3.5.9}$$

对照(3.5.8)式和(3.5.9)式,可得

$$\tilde{W}_j=\tilde{W}_j^1\oplus\tilde{W}_j^2\oplus\tilde{W}_j^3 \tag{3.5.10}$$

其中

$$\tilde{W}_j^1=V_j\otimes W_j,\tilde{W}_j^2=W_j\otimes V_j,\tilde{W}_j^3=W_j\otimes W_j \tag{3.5.11}$$

因为 $V_j\perp W_j$,所以 $\tilde{V}_j,\tilde{W}_j^1,\tilde{W}_j^2,\tilde{W}_j^3$ 必定两两正交。且 $\{\phi_{j,n}(x)\psi_{j,m}(y)\}_{n,m\in Z}$ 是 $\tilde{W}_j^1$ 的正交规范基,$\{\psi_{j,n}(x)\phi_{j,m}(y)\}_{n,m\in Z}$ 是 $\tilde{W}_j^2$ 的正交规范基,$\{\psi_{j,n}(x)\psi_{j,m}(y)\}_{n,m\in Z}$ 是 $\tilde{W}_j^3$ 的正交规范基;从而(3.5.6)式所示函数族是 $\tilde{W}_j$ 的正交规范基。

和一维情况一样,$L^2(R^2)$ 可分解为所有分辨率上小波空间的正交和

$$L^2(R^2)=\bigoplus_{j=-\infty}^{+\infty}\tilde{W}_j \tag{3.5.12}$$

因而,(3.5.7)式所示函数族是 $L^2(R^2)$ 的正交规范基。证毕。

按上述定理,分别定义 3 个离散细节信号

$$d_j^1(n,m) = \langle f(x,y), 2^{-j}\psi^1(2^{-j}x - n, 2^{-j}y - m)\rangle = \langle f(x,y), \phi_{j,n}(x)\psi_{j,m}(y)\rangle \quad (3.5.13)$$

$$d_j^2(n,m) = \langle f(x,y), 2^{-j}\psi^2(2^{-j}x - n, 2^{-j}y - m)\rangle = \langle f(x,y), \psi_{j,n}(x)\phi_{j,m}(y)\rangle \quad (3.5.14)$$

$$d_j^3(n,m) = \langle f(x,y), 2^{-j}\psi^3(2^{-j}x - n, 2^{-j}y - m)\rangle = \langle f(x,y), \psi_{j,n}(x)\psi_{j,m}(y)\rangle \quad (3.5.15)$$

信号分解的表达式如下：

$$a_{j+1}(n,m) = \sum_l \sum_k \bar{h}(2m - l)\bar{h}(2n - k)a_j(k,l) \quad (3.5.16)$$

$$d_{j+1}^1(n,m) = \sum_l \sum_k \bar{g}(2m - l)\bar{h}(2n - k)a_j(k,l) \quad (3.5.17)$$

$$d_{j+1}^2(n,m) = \sum_l \sum_k \bar{h}(2m - l)\bar{g}(2n - k)a_j(k,l) \quad (3.5.18)$$

$$d_{j+1}^3(n,m) = \sum_l \sum_k \bar{g}(2m - l)\bar{g}(2n - k)a_j(k,l) \quad (3.5.19)$$

它们是二维信号多分辨率分析的 Mallat 算法，如图 3.7 所示。

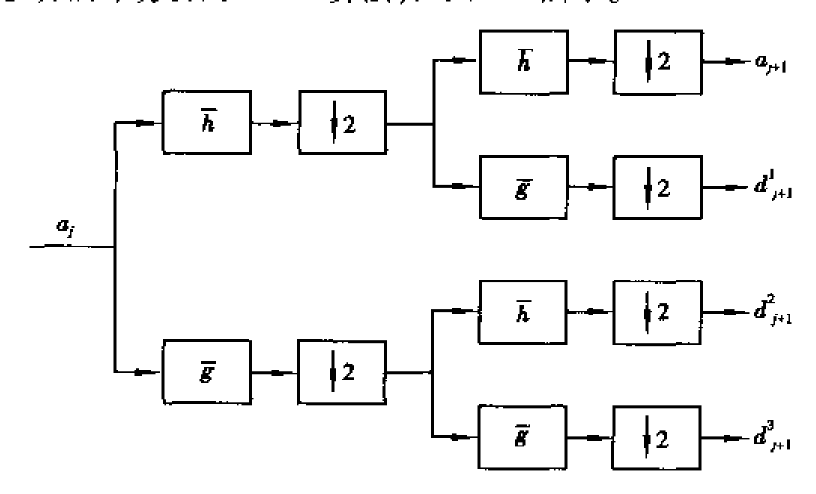


图 3.7　二维信号的分解

如原始图像为 a_0，可逐级分解为如下图像

$$\{a_J, d_j^1, d_j^2, d_j^3, 1 \leqslant j \leqslant J\} \quad (3.5.20)$$

设 a_0 有 N^2 个像素，那么 a_j 和 d_j^1, d_j^2, d_j^3 有 $2^{-j}N \times 2^{-j}N$ 个像素，每次分解都保持像素的数目不变，这是由于使用了正交基进行分解。正交小波分解最有意义的还在于它可以看成是图像在一组独立的，具有空间取向性的频率通道上的分解，a_{j+1} 对应 a_j 的低频成分，d_{j+1}^1 对应纵向高频（水平边缘），d_{j+1}^2 对应横向高频（垂直边缘），而 d_{j+1}^3 则对应两个方向的高频（斜线方向）。由图 3.7 可以看到，将 a_j 先沿横向（x 方向）滤波取偶数样本，再沿纵向（y 方向）滤波取偶数样本，便得到 4 个子图像 a_{j+1} 和 $d_{j+1}^1, d_{j+1}^2, d_{j+1}^3$。由于沿横向和纵向低通滤波得 a_{j+1}，故 a_{j+1} 对应 a_j 的低频成分；而 d_{j+1}^1 是沿横向低通滤波再沿纵向高通滤波，故 d_{j+1}^1 对应纵向高频（水平边缘）；d_{j+1}^2 是沿横向高通滤波再沿纵向低通滤波，故 d_{j+1}^2 对应横向高频（垂直边缘）；d_{j+1}^3 沿横向和纵向都是高通滤波，故 d_{j+1}^3 则对应两个方向的高频（斜线方向）。这就解释了小波多分辨率分析确实是对图像在不同方向上频率成分的分解。由于采用了正交小波基分解，4 个子图像是独立的，不存在信息冗余，正因为这样，才使每次分解都保持像素的数目不变。

二维信号的 Mallat 重构算法如下：

$$a_{j+1}(n,m) = \sum_k \sum_l h(n - 2k)h(m - 2l)a_j(k,l) + \sum_k \sum_l g(n - 2k)h(m - 2l)d_j^1(k,l) +$$

$$\sum_{k}\sum_{l} h(n-2k)g(m-2l)d_j^2(k,l) + \sum_{k}\sum_{l} g(n-2k)g(m-2l)d_j^3(k,l) \tag{3.5.21}$$

上述算法表示为图 3.8。

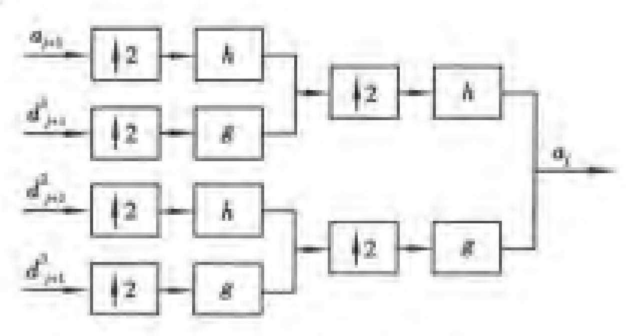


图 3.8　二维信号的重构

对图 3.7 及图 3.8 所示的二维小波分解与重构算法，利用其可分离特性，在算法实现时分别按行进行一维小波变换，然后对按行变换后的数据再按列进行一维小波变换来完成。与一维的情形类似，在实际应用中，由于图像信号总是有限区域的，也存在如何处理边界的问题。典型的处理方法是周期延拓和对称延拓。在用小波变换进行图像压缩时，由于边界的不连续性，会使得边界附近的小波变换系数的衰减变慢，从而影响图像的压缩比，因而在图像压缩应用中，若使用的是具有对称性质的双正交小波滤波器，一般对边界采用对称延拓的方式，使边界保持连续，以提高压缩性能。

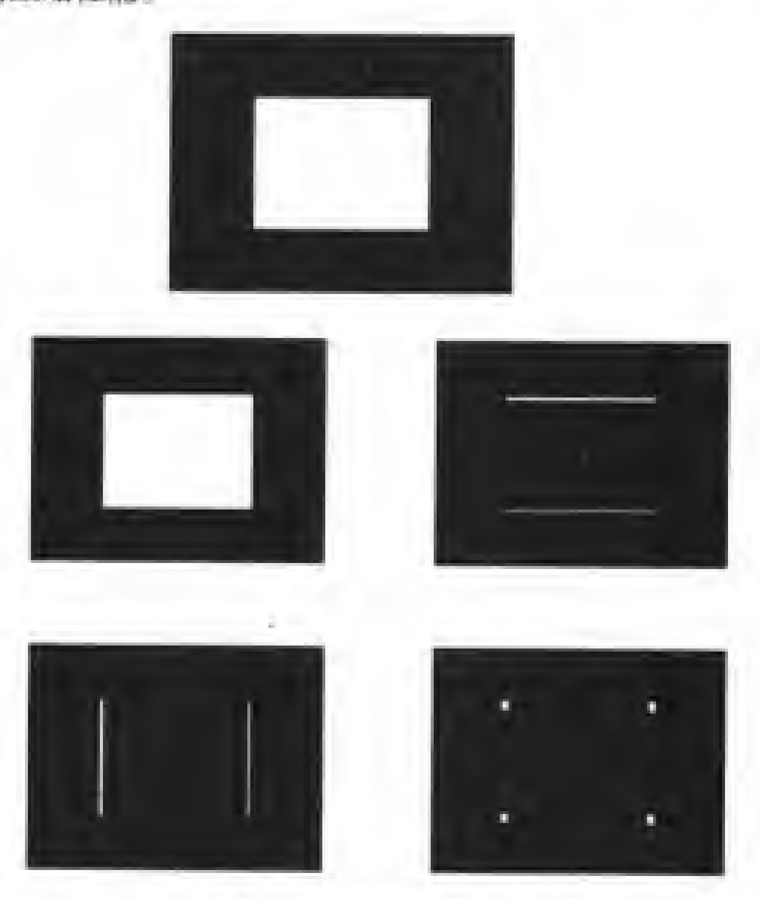


图 3.9　图像信号的 Mallat 分解

在 Matlab 小波工具箱中，函数 dwt2 用来作 1 级二维离散小波变换。图 3.9 上方为原始图

像，它是黑色背景中的一个白方块。用 dwt2 将它分解成 4 个子图像，一般将 a_{j+1} 和 d_{j+1}^1，d_{j+1}^2，d_{j+1}^3 分别置于左上角和右上角，左下角，右下角，如图 3.8 下方所示。图 3.8 最清楚不过地说明了二维多分辨率分析的空间取向性：d_{j+1}^1 确实检测出两条水平边缘，d_{j+1}^2 确实检测出两条垂直边缘，d_{j+1}^3 确实检测出 4 个边角。在上面的分解过程中，选用了 db2 小波，如改换为其他小波，会得到稍微不同的结果。但如选用更光滑，从而支撑宽度更大的 db20 小波，结果就很不理想了。一般来说，对比较奇异的图像，如图 3.9 上方所示的图像，应选用支撑宽度较小的小波，而对较光滑的图像，则应选用支撑宽度较大的小波。

3.6 图像小波分析

小波分析在图像处理中的应用是一个十分重要的研究领域，目前已取得了许多显著的研究及应用成果。在用 Matlab 小波工具箱进行图像处理时，经常要调用图像处理工具箱和信号处理工具箱的函数和命令，同时要特别注意 Matlab 所支持的图像类型和数据类型，才能达到预期的效果。

Matlab 支持 4 种类型的图像，它们分别是 RGB 真彩色图像、索引图像、灰度图像和二值图像。不同类型的图像实际上是颜色和亮度的不同编码方式。无论什么类型的图像，都表示为数组。Matlab 支持的图像数组元素的数据类型有 3 种：双精度浮点数、16 位整数和 8 位整数。

RGB 图像表示为三维数组 $m\times n\times 3$，实际上是 3 个 m 行 $\times n$ 列的二维数组，分别表示红、绿、蓝三分色的亮度。例如像素(8,5)的三分色的亮度分别存储在(8,5,1)、(8,5,2)和(8,5,3)3个数组元素中。RGB 图像的数据类型可以是双精度、16 位整型和 8 位整型。双精度 RGB 图像数组元素的取值范围在 0 与 1 之间。如数据类型是 8 位整型，颜色深度为 24 比特，可表示的颜色多达 16×10^6 种，已非常接近真实世界的色彩，故 RGB 图像又称真彩色图像。

索引图像包括图像数组与颜色图数组，图像数组是 m 行 $\times n$ 列的二维数组，数组元素的值作为像素颜色的索引，颜色图数组是 c 行 $\times 3$ 列数组，索引指向颜色图数组的行，查找出像素红、绿、蓝三分色的亮度，不难理解，图像数组元素的值应该是整数。显然，c 值的大小表示了图像颜色的深度，例如 $c=1\ 000$，表示图像有 1 000 种不同的颜色。颜色图数组的数据类型只能是双精度型，数组元素的取值范围在 0 与 1 之间，颜色图数组的每一行指定出某一单色的红、绿、蓝三分量的亮度。图像数组的数据类型则可以是双精度、16 位整型和 8 位整型。如图像数组的数据类型是双精度型，数组元素的值直接作为查找像素颜色的索引，例如图像数组元素的值为 1，查找颜色图数组的第 1 行，为 2 则查找颜色图数组的第 2 行，依此类推。如图像数组的数据类型是 16 位整型或 8 位整型，这时图像数组元素的值可能等于 0，而颜色图数组的起始行为第 1 行，所以必须加上一个等于 1 的偏移量，即图像数组元素的值加 1 作为查找像素颜色的索引。例如图像数组元素的值为 0，查找颜色图数组的第 1 行，为 1 则查找颜色图数组的第 2 行，依此类推。

灰度图像(或称强度图像)表示为 m 行 $\times n$ 列的二维数组，数组元素的值就是相应像素的灰度等级。灰度图像的数据类型可以是双精度、16 位整型和 8 位整型。数组元素的值为 0 表示相应像素为黑，数组元素的值为 1，256 或 65 535 时，表示白。Matlab 在存储灰度图像时没有颜色图，但在显示时有颜色图。

二值图像可以说是一种特殊类型的灰度图像，它只有黑白两个灰度等级。二值图像的数据类型只能是双精度或8位整型，不可能是16位整型，显然用8位整型更节省存储空间。

Matlab 提供了若干实现图像类型和色度空间互相转换的函数。

数字图像在计算机中是以数据文件的形式存储的，Matlab 支持的图像文件格式包括：BMP（Windows Bitmap）、HDF（Hierarchical Data Format）、JPEG（Joint Photographical Experts Group）、PCX（Windows Paintbrush）、TIFF（Tagged Image File Format）、XWD（X Windows Dump）。

对于彩色图像，Matlab 小波工具箱只支持索引图像，所以 RGB 图像在用小波工具箱函数处理之前必须转换成索引图像。如颜色图是平滑的，则小波分解可直接用于索引图像；否则，可以将它转换成颜色图平滑的灰度图像。图像数组元素的值应该是整数，但也容许是非整数。在进行数值计算的过程中，图像数组元素的值为非整数不会带来任何问题，但在显示图像时，小于1的数都视为1，大于 c 的数都视为 c（c 为颜色图数组的行数），在1和 c 之间的非整数四舍五入为最接近的整数。此外，小波分解产生的系数（即离散逼近信号和3个方向的离散细节信号）、逼近和细节（经过插值的逼近和细节信号）不再是索引数组。

例 3.7　将 Matlab 小波工具箱中存储的图像用 wvdec2 函数作两级分解。程序如下：

```
clear all
% 装入并显示索引图像
load wbarb;image(X); colormap(map); colorbar;
% 将索引图像转化成灰度图像并显示
I = ind2gray(X,map);figure; I = imadjust(I,stretchlim(I),[0,1]);imshow(I);
% 做二维小波分解并显示其结果
wname = 'sym2';
[C,S] = wavedec2(I,2,wname);
cA1 = appcoef2(C,S,wname,1);
[cH1,cV1,cD1] = detcoef2('all',C,S,1);
cA2 = appcoef2(C,S,wname,2);
[cH2,cV2,cD2] = detcoef2('all',C,S,2);
A1 = wrcoef2('a',C,S,wname,1);
H1 = wrcoef2('h',C,S,wname,1);
V1 = wrcoef2('v',C,S,wname,1);
D1 = wrcoef2('d',C,S,wname,1);
A2 = wrcoef2('a',C,S,wname,2);
H2 = wrcoef2('h',C,S,wname,2);
V2 = wrcoef2('v',C,S,wname,2);
D2 = wrcoef2('d',C,S,wname,2);
figure;
subplot(2,2,1);A1 = abs(A1);A1 = imadjust(A1,stretchlim(A1),[0,1]);imshow(A1);
title('Approximation A1')
subplot(2,2,2);H1 = abs(H1);H1 = imadjust(H1,stretchlim(H1),[0,1]);imshow(H1);
title('Horizontal Detail H1')
```

```
subplot(2,2,3);V1 = abs(V1);V1 = imadjust(V1,stretchlim(V1),[0,1]);imshow(V1);
title('Vertical Detail V1')
subplot(2,2,4);D1 = abs(D1);D1 = imadjust(D1,stretchlim(D1),[0,1]);imshow(D1);
title('Diagonal Detail D1')
figure;
subplot(2,2,1);A2 = abs(A2);A2 = imadjust(A2,stretchlim(A2),[0,1]);imshow(A2);
title('Approximation A2')
subplot(2,2,2);H2 = abs(H2);H2 = imadjust(H2,stretchlim(H2),[0,1]);imshow(H2);
title('Horizontal Detail H2')
subplot(2,2,3);V2 = abs(V2);V2 = imadjust(V2,stretchlim(V2),[0,1]);imshow(V2);
title('Vertical Detail V2')
subplot(2,2,4);D2 = abs(D2);D2 = imadjust(D2,stretchlim(D2),[0,1]);imshow(D2);
title('Diagonal Detail D2')
```

原始图像是一个索引图像，其文件名为 wbarb，图像数组与颜色图数组名分别为 X 和 map，程序运行之后，可以看到颜色图是平滑的。将它转换成灰度图像 I 之后作为小波分解的原始图像 A0，做了两级分解之后第 1 级的逼近及横向、纵向、斜向细节为 A1 和 H1、V1、D1，第 2 级的逼近及横向、纵向、斜向细节为 A2 和 H2、V2、D2，如图 3.10 所示。



A0

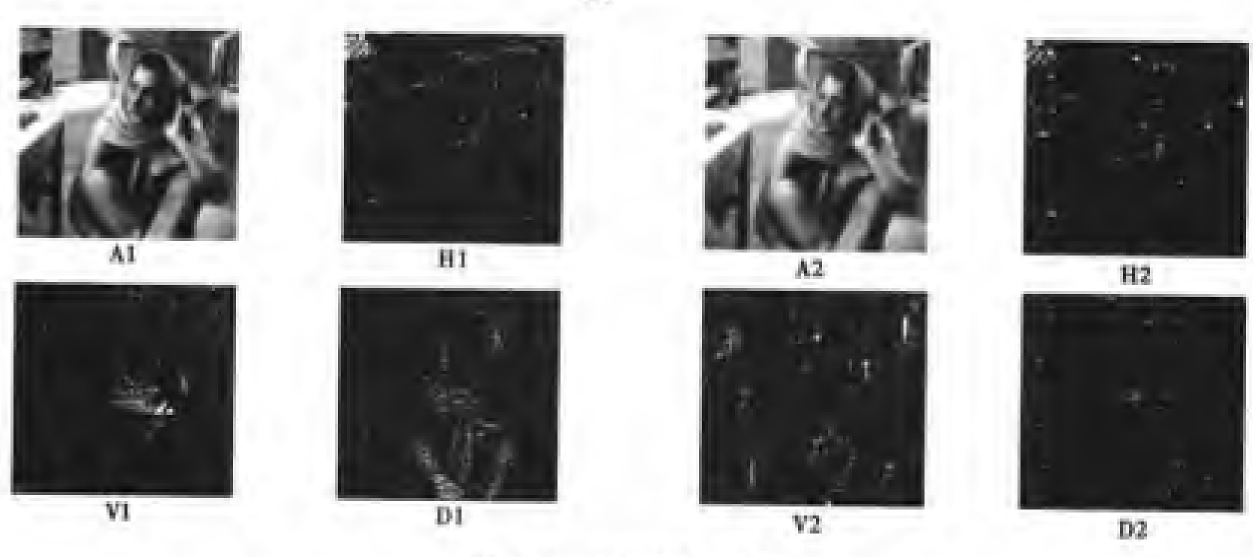


图 3.10 图像的小波分解

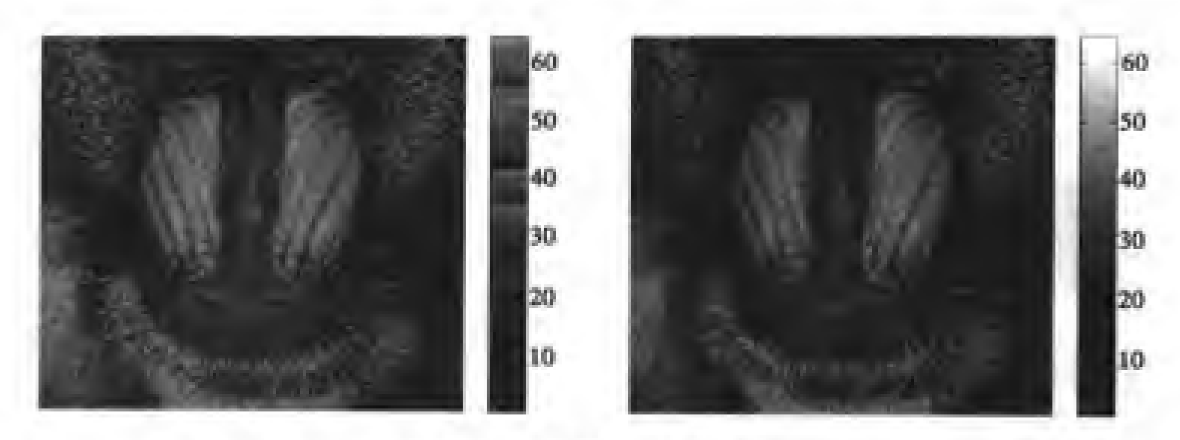


图3.11 索引图像颜色图平滑

例3.8 将索引图像的颜色图平滑化。程序如下：

```
% 装入文件名为"xpmndrll"的索引图像,并显示
load xpmndrll
image(X2);colormap(map); colorbar;title('Original Color Indexed Image')
% 转换成颜色图平滑的灰度索引图像,并显示
I = ind2gray(X2,map);
n = size(map,1);
X = round(I * (n - 1)) + 1;
map2 = gray(n);
figure;image(X);colormap(map2), colorbar;title('Processed Gray Scale Indexed Image')
% 将结果存为文件名为"baboon"的磁盘文件,其图像数组名是"baboon",颜色图数组名是"map"
baboon = X;map = map2;save baboon baboon map
```

如图3.11左图所示，索引图像的颜色图不平滑；经过处理之后转换成颜色图平滑的灰度图像。

例3.9 如果已有一幅数码相机拍摄的真彩色JPEG图像存入Matlab，其文件名为'example'，可通过如下程序将它转换成适当的索引图像：

```
% 调用 imfinfo 函数显示数码相机拍摄的真彩色 JPEG 图像的相关性质
Info = imfinfo('example.jpg','jpg')
% 读入图像并转换成索引图像后加以显示
RGB1 = imread('example.jpg');
[X1,map1] = rgb2ind(RGB1,128);
image(X1);colormap(map1);colorbar;
% 修改图像数组大小,并将其数据类型由8位整型修改为双精度浮点型 %
X2 = imresize(X1,0.5);X3 = double(X2) + 1;
figure;image(X3);colormap(map1);colorbar;
```

第 4 章

正交小波的构造

第 3 章指出可以从尺度滤波器生成多分辨率分析从而构造正交小波，Daubechies 正是通过设计尺度滤波器构造出几种具有特定性质的小波，在小波分析中获得了广泛的应用。了解小波的构造对理解它们的性质，从而在实际应用中正确地选择小波是大有裨益的。

4.1 尺度滤波器的性质

多分辨率分析中，尺度函数 $\phi(t)$ 和尺度滤波器 $h(n)$ 满足如下二尺度方程

$$\phi(t) = \sqrt{2}\sum_{n} h(n)\phi(2t-n) \tag{4.1.1}$$

这是一个十分重要的关系式，它意味着可以通过设计 $h(n)$ 得到尺度函数 $\phi(t)$，再根据 (3.2.10) 式计算 $g(n)$ 从而构造正交小波 $\psi(t)$。

多分辨率分析中的滤波器 $g(n)$ 由式

$$g(n) = \pm(-1)^{1-n}h(1-n) \tag{4.1.2}$$

计算。在信号处理中，称 $h(n)$ 和 $g(n)$ 为正交镜像滤波器 QMF (quadrature mirror filter)，或共轭镜像滤波器 CMF (conjugate mirror filter)。

二尺度方程(4.1.1)蕴涵了这样一个事实：$\{\phi_{j,n}(t)\}_{n\in Z}$ 是 V_j 的正交规范基。而 $\phi(t)$ 二进伸缩和平移构成正交规范基的充分必要条件是

$$\sum_{k} |\hat{\phi}(\omega+2k\pi)|^2 = 1$$

对(4.1.1)式两边作傅立叶变换得

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}}H\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right)$$

将它代入正交规范基的充分必要条件，有

$$\sum_{k} |H(\omega/2+k\pi)|^2 |\hat{\phi}(\omega/2+k\pi)|^2 = 2$$

在上式中将 k 分为奇偶两组分别求和，并考虑到 $H(\omega)$ 是 2π 的周期函数，得

$$|H(\omega/2)|^2 + |H(\omega/2+\pi)|^2 = 2$$

将 $\omega/2$ 换成 ω，有

$$|H(\omega)|^2 + |H(\omega+\pi)|^2 = 2 \tag{4.1.3}$$

它实际上是 $\phi(t)$ 正交性的等价表达式。在构造小波的过程中，上式将起关键作用。

类似的，$\{\psi_{j,n}(t)\}_{n\in Z}$ 是 W_j 的正交规范基的等价表达式为

$$|G(\omega)|^2 + |G(\omega+\pi)|^2 = 2 \tag{4.1.4}$$

还可以证明[9]：$\phi(t)$ 与 $\psi(t)$ 的平移正交性等价于

$$H(\omega)G^*(\omega) + H(\omega+\pi)G^*(\omega+\pi) = 0 \tag{4.1.5}$$

上列 3 式指出共轭镜像滤波器具有一定的正交性。

定理 4.1　共轭镜像滤波器的正交性质的时域表达形式是：

$$\sum_n h(n)h(n-2k) = \delta_k \tag{4.1.6}$$

$$\sum_n g(n)g(n-2k) = \delta_k \tag{4.1.7}$$

$$\sum_n h(n)g(n-2k) = 0 \tag{4.1.8}$$

证明　由二尺度方程(4.1.1)，有

$$\begin{aligned}\phi(t-k) &= \sqrt{2}\sum_n h(n)\phi(2t-2k-n)\\ &= \sqrt{2}\sum_l h(l-2k)\phi(2t-l)\end{aligned}$$

所以，

$$\begin{aligned}\delta_k &= \langle \phi(t), \phi(t-k)\rangle\\ &= 2\langle \sum_n h(n)\phi(2t-n), \sum_l h(l-2k)\phi(2t-l)\rangle\\ &= 2\sum_n \sum_l h(n)h(l-2k)\langle \phi(2t-n), \phi(2t-l)\rangle\end{aligned}$$

由 $\phi(t)$ 的平移正交性，知 $2\langle\phi(2t-n),\phi(2t-l)\rangle = \delta_{n,l}$，故

$$\begin{aligned}\delta_k &= \langle \phi(t), \phi(t-k)\rangle\\ &= \sum_n \sum_l h(n)h(l-2k)\delta_{n,l}\\ &= \sum_n h(n)h(n-2k)\end{aligned}$$

同理，由二尺度方程(3.3.12)及 $\psi(t)$ 的平移正交性可证(4.1.7)式。而证明(4.1.8)式则要用到 $\phi(t)$ 和 $\psi(t)$ 之间的正交性。证毕。

(4.1.6)式说明 $h(n)$ 与其偶数平移是正交的。由证明过程知，(4.1.6)式是 $\phi(t)$ 的平移正交性的等价式，(4.1.7)式是 $\psi(t)$ 的平移正交性的等价式，而(4.1.8)式是 $\phi(t)$ 与 $\psi(t)$ 正交性的等价式。

二尺度方程(4.1.1)的频域等价表达式为：$\sqrt{2}\hat{\phi}(\omega) = H(\omega/2)\hat{\phi}(\omega/2)$，令 $\omega=0$，由于 $\hat{\phi}(0)\neq 0$，所以有

$$H(0) = \sqrt{2} \tag{4.1.9}$$

它意味着 $h(n)$ 的系数必须满足

$$\sum_n h(n) = \sqrt{2} \tag{4.1.10}$$

这是 $h(n)$ 必须满足的最弱条件。

令(4.1.6)式中 $k=0$,有

$$\sum_n |h(n)^2| = 1 \tag{4.1.11}$$

即尺度滤波器系数的范数等于1。

此外,$h(n)$ 的偶数系数之和应等于其奇数系数之和,即

$$\sum_n h(2n) = \sum_n h(2n+1) = \frac{1}{\sqrt{2}} \tag{4.1.12}$$

证明 令(4.1.3)式中 $\omega=\pi$,有

$$H(\pi) = 0 \tag{4.1.13}$$

根据傅立叶变换的定义,得

$$\begin{aligned} H(\pi) &= \sum_n h(n)\mathrm{e}^{-in\pi} = \sum_n (-1)^n h(n) \\ &= \sum_n h(2n) - \sum_n h(2n+1) = 0 \end{aligned}$$

再考虑到(4.1.10)式,便可由上式得到(4.1.12)式。证毕。

(4.1.12)式比(4.1.10)式强,但比正交性条件(4.1.6)式弱。假若 $h(n)$ 是一个长度为 N 的 FIR 滤波器,则由(4.1.10)式和(4.1.6)式所规定的约束条件可以列出 $N/2+1$ 个方程,从而求出 $h(n)$ 的系数。但为了得到性质更好的小波,应该利用剩余的自由度对 $h(n)$ 施加更多的约束条件,其中最重要的就是小波消失矩。

4.2 小波基的性质

小波基的应用大多是充分发挥其用很少的非零小波系数去有效逼近特殊函数类的能力。图像压缩是这样,噪声去除和快速计算也是这样。故 $\psi(t)$ 的设计必须被最优化以产生尽可能多的小幅值小波系数。这个性质主要依赖于信号的正则性,小波的消失矩阶数及其支集的大小。为了由 $h(n)$ 构造合适的小波,应该将小波消失矩阶数及其支集大小与 $h(n)$ 联系起来。

4.2.1 小波消失矩及其支集

1.5 节曾论述过消失矩,如

$$\int_{-\infty}^{+\infty} t^k \psi(t)\mathrm{d}t = 0, \quad k = 0,1,\cdots,K-1 \tag{4.2.1}$$

则 $\psi(t)$ 具有 K 阶消失矩。这意味着 $\psi(t)$ 与任何小于 K 次的多项式是正交的。如果 $f(t)$ 是局部 C^k 的,那么在一个小区间上它可以用 $k-1$ 阶泰勒多项式很好地逼近。若 $k<K$,则小波与这些泰勒多项式正交,因而对细尺度产生小幅值小波系数。下述命题将 $\psi(t)$ 的消失矩阶数与 $\hat{\psi}(\omega)$ 在 $\omega=0$ 处导数为零的阶数,以及 $H(\omega)$ 在 $\omega=\pi$ 处的零点重数联系起来。它同时表明直到 $K-1$ 次的多项式均可表示成 $\{\phi(t-n)\}_{n\in Z}$ 的线性组合。

定理4.2　（消失矩）

设$\psi(t)$和$\phi(t)$是生成一组正交基的小波和尺度函数，且当$t\to\pm\infty$时有足够快的衰减，即$|\psi(t)|=O((1+t^2)^{-K/2-1})$且$|\phi(t)|=O((1+t^2)^{-K/2-1})$，则下面4个命题等价：

(1)小波$\psi(t)$有K阶消失矩；

(2)$\hat{\psi}(\omega)$及其前$K-1$阶导数在$\omega=0$处为零；

(3)$H(\omega)$及其前$K-1$阶导数在$\omega=\pi$处为零；

(4)任何$K-1$次多项式可以表示成$\{\phi(t-n)\}_{n\in Z}$的线性组合。

由$|\phi(t)|$和$|\psi(t)|$的衰减性知$\hat{\phi}(\omega)$和$\hat{\psi}(\omega)$是K次连续可微的。根据傅立叶变换的性质，(4.2.1)式等价于：

$$\frac{\mathrm{d}^k}{\mathrm{d}\omega^k}\hat{\psi}(\omega)\bigg|_{\omega=0}=0\quad k=0,1,\cdots,K-1 \tag{4.2.2}$$

由(3.2.10)式，有

$$\sqrt{2}\hat{\psi}(2\omega)=G(\omega)\hat{\phi}(\omega)$$

由于$\hat{\phi}(0)\neq0$，$G(0)=0$，所以(4.2.2)式意味着$G(\omega)$在$\omega=0$处的前$K-1$阶导数必然为零。又因为

$$G(\omega)=\mathrm{e}^{-\mathrm{i}\omega}H^*(\omega+\pi)$$

所以，$H(\omega)$及其前$K-1$阶导数在$\omega=\pi$处为零。如果$h(n)$是一个实系数FIR滤波器，则$H(\omega)$在$\omega=\pi$有K重零点，从而$H(\omega)$可写成

$$H(\omega)=\sqrt{2}\left(\frac{1+\mathrm{e}^{-\mathrm{i}\omega}}{2}\right)^K Q(\omega) \tag{4.2.3}$$

其中，$Q(\omega)$是一个实系数多项式，且在$\omega=\pi$处既无极点又无零点。

4.2.2　小波支集

如果$f(t)$在t_0处有一孤立奇点，且t_0包含在$\psi_{j,n}$的支集内，那么小波系数$\langle f,\psi_{j,n}\rangle$一般有大的幅值。如果$\psi(t)$有一长度为$M$的紧支集，则对每个尺度$2^j$，均有$M$个$\{\psi_{j,n}\}_{n\in Z}$的支集包含$t_0$，从而产生大幅值的小波系数。为了使大幅值的小波系数最少，必须减小$\psi(t)$的支集长度。下述定理将$h(n)$的支集长度和$\phi(t)$与$\psi(t)$的支集联系起来。

定理4.3　（紧支集）

尺度函数$\phi(t)$有紧支集当且仅当$h(n)$有紧支集，且两者的支集相等。如果$\phi(t)$和$h(n)$的支集是$[N_1,N_2]$，那么，$\psi(t)$的支集是$[(N_1-N_2+1)/2,(N_2-N_1+1)/2]$。

这个定理很容易由二尺度方程加以证明。由上述定理可以看到，$\psi(t)$的支集长度为N_2-N_1，支集中心位于1/2。

一般说来，一个函数的支集长度与其消失矩阶数是独立的。但对正交小波来说，具有K阶消失矩意味着其支集长度至少是$2K-1$。所以在选择正交小波时，必须在支集长度和消失矩阶数之间折中。如果信号有很少的孤立奇异点，在奇异点之间很光滑，应选择具有高阶消失矩的小波以使大量的小波系数幅值很小。而当奇异点靠得很近时，最好选择支集长度较小的小波，以便于区分这些奇异点，同时也可以减少大幅值的小波系数。

4.2.3 正则性

在 1.5 节曾指出，如函数$f(t)$在点 t_0 的邻域$[t_0-h,t_0+h]$是 m 次可微的，则在该邻域内的泰勒多项式

$$P(t)=\sum_{k=0}^{m-1}\frac{f^{(k)}(t_0)}{k!}(t-t_0)^k \tag{4.2.4}$$

是对$f(t)$的一个很好的逼近，逼近误差满足

$$|f(t)-P(t)|\leqslant\frac{|t-t_0|^m}{m!}\sup_{u\in[t_0-h,t_0+h]}|f^{(m)}(u)| \tag{4.2.5}$$

为了更精细的定量描述函数的正则性，首先引入 Lipschitz 指数 α，α 不必是一个正整数。在第 7 章将进一步看到，α 可以小于 1 甚至是一个负数，这样就可以用 Lipschitz 指数来描述函数的奇异性。

定义 4.1 对函数$f(t)$，如存在常数 $C>0$，以及 $m=\lfloor\alpha\rfloor$阶多项式 $P(t)$，使得

$$\forall t\in R,|f(t)-P(t)|\leqslant C|t-t_0|^{\alpha} \tag{4.2.6}$$

则称函数$f(t)$在点 t_0Lipschitz α。$m=\lfloor\alpha\rfloor$表示 m 为小于 α 的最大整数。如对所有的 $t_0\in[a,b]$以及一个与 t_0 无关的常数 C，使得(4.2.6)式成立，则称函数$f(t)$在区间$[a,b]$一致 Lipschitz α。使(4.2.6)式成立的 α 的上界定义为 Lipschitz 正则性。

当小波分析用于数据压缩和去噪时，要量化小波系数并做阈值处理，然后在进行信号重构时，量化和阈值处理所带来的误差 ε 会使每个分量产生误差 $\varepsilon\psi_{j,n}$，从而引起重构信号的降质。如小波比较平滑，则误差 $\varepsilon\psi_{j,n}$也比较平滑。人眼对平滑误差比较不敏感，所以在图像压缩中，应尽量选用正则性较好的小波作图像重构。下述定理将尺度函数和小波的一致 Lipschitz 正则性与 $H(\omega)$在 $\omega=\pi$ 处的零点重数联系起来。这时 $H(\omega)$仍然可以写成(4.2.3)式所示的因式分解形式。

定理 4.4 如 $\sup_{\omega\in R}|Q(\omega)|=B$，则 ψ 和 ϕ 的一致 Lipschitz α 满足

$$\alpha<\alpha_0=K-\log_2 B-1 \tag{4.2.7}$$

由上述定理可以看到：如 $B<2^{K-1}$，则 $\alpha_0>0$，这意味着 ψ 和 ϕ 至少是一致连续的。如对任意的 $m>0$，$B<2^{K-1-m}$，则 $\alpha_0>m$，这时 ψ 和 ϕ 是 m 次连续可微的。虽然不能断定高阶消失矩的小波具有更好的 Lipschitz 正则性，因为随着消失矩阶数 K 增加，B 也可能增加；但对 Daubechies 小波和样条小波这样一些重要的小波而言，B 的增加要比 K 的增加慢得多，所以高阶消失矩的小波具有更好的 Lipschitz 正则性。

4.3 Daubechies 小波

著名数学家 Daubechies 根据尺度滤波器的性质，指出了在频率域构造小波的途径，为小波理论的发展作出了卓越的贡献。

考虑支集为$[0,N-1]$的实系数因果滤波器 $h(n)$，这时 $\psi(t)$是支集为$[1-N/2,N/2]$的紧支小波。Daubechies 小波是在给定消失矩的条件下，具有最小支集的小波。

对于支集为$[0,N-1]$的实系数因果滤波器 $h(n)$，其傅立叶变换为如下三角多项式：

$$H(\omega)=\sum_{n=0}^{N-1}h(n)e^{-i\omega n} \tag{4.3.1}$$

小波有 K 阶消失矩意味着 $H(\omega)$ 在 $\omega=\pi$ 处有 K 重零点，或者说 $H(z)$ 在 $z=-1$ 处有 K 重零点，所以 $h(n)$ 的傅立叶变换可写成

$$H(\omega)=\sqrt{2}\left(\frac{1+e^{-i\omega}}{2}\right)^K Q(\omega) \tag{4.3.2}$$

而其 Z 变换为如下多项式：

$$H(z)=\sum_{n=0}^{N-1}h(n)z^{-n}=\sqrt{2}\left(\frac{1+z^{-1}}{2}\right)^K Q(z) \tag{4.3.3}$$

其中，$Q(z)$ 是 z^{-1} 的 $N-1-K$ 次多项式。由于 $h(n)$ 的系数是实的，所以 $|H(\omega)|^2$ 是偶函数，可写成 $\cos\omega$ 的多项式，从而 $|Q(\omega)|^2$ 也可写成 $\cos\omega$ 的多项式，并进而写成 $\sin^2(\omega/2)$ 的多项式：

$$\begin{aligned}|Q(\omega)|^2&=|Q(z)Q(z^{-1})|_{z=e^{i\omega}}\\&=P[\sin^2(\omega/2)]\end{aligned} \tag{4.3.4}$$

其中，P 是一个 $N-1-K$ 次多项式。令 $y=sin^2(\omega/2)$，则由(4.3.2)式得

$$\begin{aligned}|H(\omega)|^2&=2\cos^{2K}(\omega/2)\,|Q(\omega)|^2\\&=2(1-y)^KP(y)\end{aligned} \tag{4.3.5}$$

考虑到尺度滤波器正交性的频域表达式：

$$|H(\omega)|^2+|H(\omega+\pi)|^2=2 \tag{4.3.6}$$

由上列两式可得

$$y^KP(1-y)+(1-y)^KP(y)=1 \tag{4.3.7}$$

显然，方程(4.3.7)应满足如下约束条件：

$$P(y)\geqslant 0,y\in[0,1] \tag{4.3.8}$$

为了得到 $h(n)$ 的系数，必须求解关于多项式 P 的方程(4.3.7)。关于它，有如下 Bezout 定理。

定理 4.5 (Bezout 定理)

设 $Q_1(y)$ 和 $Q_2(y)$ 是 n_1 和 n_2 次多项式，它们没有公共零点。则存在次数分别为 n_2-1 和 n_1-1 的两个惟一多项式 $P_1(y)$ 和 $P_2(y)$，使得

$$P_1(y)Q_1(y)+P_2(y)Q_2(y)=1 \tag{4.3.9}$$

将 Bezout 定理用于方程(4.3.7)：令 $Q_1(y)=y^k$，$Q_2(y)=(1-y)^K$ 它们显然没有公共零点。于是存在两个惟一的 $K-1$ 次多项式 $P_1(y)$ 和 $P_2(y)$，使得

$$y^KP_1(y)+(1-y)^KP_2(y)=1$$

令 $x=1-y$，则有

$$(1-x)^KP_1(1-x)+x^KP_2(1-x)=1$$

对照上列两式知 $P_1(y)=P_2(1-y)=P(1-y)$，所以方程(4.3.7)存在一个惟一的多项式解 $P(y)$。

可以证明[8]，方程(4.3.7)的通解是

$$P(y)=\sum_{K=0}^{K-1}\binom{K-1+k}{k}y^k+y^KR\left(\frac{1}{2}-y\right) \tag{4.3.10}$$

其中，$R(y)$ 是一奇次多项式，使得 $P(y)\geqslant 0$，$\forall y\in[0,1]$。

当 $R(y)=0$，$P(y)$ 是一个 $K-1$ 次多项式，故 $N-1-K=K-1$，即 $N=2K$。如 $R(y)\neq 0$，$P(y)$ 是一个高于 $K-1$ 次的多项式，应有 $N-1-K>K-1$，即 $N>2K$。

构造 Daubechies 小波时，令 $R(y)=0$，此时尺度滤波器 $h(n)$ 具有最小长度 $N=2K$。当 $R(y)\neq 0$时，可以附加小波消失矩以外的其他约束条件，构造具有更好性质的小波，但这些小波将具有更大的支集长度。所以 Daubechies 小波也是在给定滤波器长度的条件下，具有最大消失矩的小波。对于 K 阶 Daubechies 小波，它具有 K 阶消失矩，$\phi(t)$ 和 $h(n)$ 的支集是 $[0,2K-1]$，$\psi(t)$ 的支集是 $[1-K,K]$；也就是说，$\psi(t)$ 的支集长度是 $2K-1$，而 $h(n)$ 的支集长度是 $2K$。

对于 Daubechies 小波

$$P(y)=P_K(y)=\sum_{k=0}^{K-1}\binom{K-1+k}{k}y^k,K=\frac{N}{2} \tag{4.3.11}$$

这时，$P(y)$ 是一个 $K-1$ 次多项式，其系数为

$$p_k=\binom{K-1+k}{k},k=0,1,\cdots,K-1 \tag{4.3.12}$$

也可由下列迭代计算式

$$p_k=\sum_{l=0}^{k-1}(-1)^{k-l+1}\binom{K}{k-l}p_l,\quad k=1,2,\cdots,K-1 \tag{4.3.13}$$

来计算 $P(y)$ 的系数。然后将 $P\left[\sin^2(\omega/2)\right]$ 代入(4.3.4)式便可算出 $Q(z)$，再由(4.3.3)式计算出尺度滤波器的系数。由于 $Q(z)$ 和 $Q(z^{-1})$ 的零点互为倒数，如选取 $Q(z)$ 的所有零点均位于单位圆内，那么，尺度滤波器将具有最小相位。这样得到的因果滤波器 $h(n)$ 的能量主要集中在 $n=0$ 附近。这意味着 $h(n)$ 很不对称，不具有线性相位，因此，$\phi(t)$ 和 $\psi(t)$ 也很不对称。

例 4.1 取 $K=2$，则 $p_0=1$，$p_1=2$，由(4.3.4)式得

$$|Q(\omega)|^2=P\left(\sin^2(\omega/2)\right)=1+2\sin^2(\omega/2)=2-\cos\omega=2-\frac{1}{2}(e^{i\omega}+e^{-i\omega})$$

$$Q(z)Q(z^{-1})=2-\frac{1}{2}(z+z^{-1})$$

令上式为零求得 $Q(z)Q(z^{-1})$ 的零点为 $z_{1,2}=2\pm\sqrt{3}$，取 $2+\sqrt{3}$ 得

$$Q(z^{-1})=c\left(z-(2+\sqrt{3})\right)$$

$$Q(z)=c\left(z^{-1}-(2+\sqrt{3})\right)$$

之所以要取 $2+\sqrt{3}$，是考虑到 $Q(z)$ 的零点位于单位圆内。上列两式中，c 是一待定系数，可由下式确定

$$Q(z)Q(z^{-1})=c^2\left(z^{-1}-(2+\sqrt{3})\right)\left(z-(2+\sqrt{3})\right)=2-\frac{1}{2}(z+z^{-1})$$

可以通过比较同类项的系数来确定 c，也可以简单地令 $z=1$ 由上式求得 $c=(1-\sqrt{3})/2$。将

$$Q(z)=\frac{1-\sqrt{3}}{2}\left(z^{-1}-(2+\sqrt{3})\right)$$

代入(4.3.3)式便可求得尺度滤波器系数如下：

n	0	1	2	3
$h(n)$	$\sqrt{2}(1+\sqrt{3})/8$	$\sqrt{2}(3+\sqrt{3})/8$	$\sqrt{2}(3-\sqrt{3})/8$	$\sqrt{2}(1-\sqrt{3})/8$

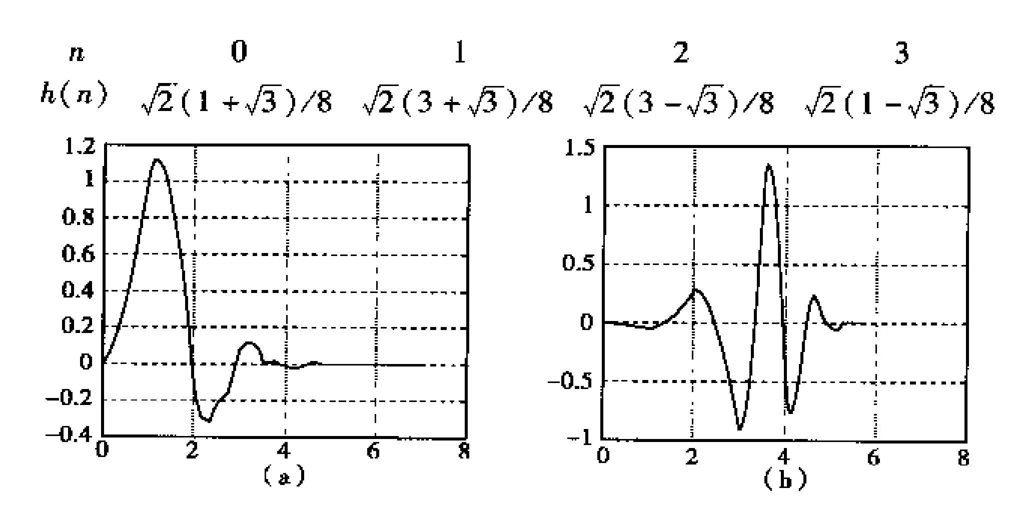


图4.1　db4 尺度函数(a)及小波(b)

按(4.1.10)和(4.1.11)式,尺度滤波器系数之和为$\sqrt{2}$,而其范数为1。如改变规范化的方式,将上述尺度滤波器系数均除以$\sqrt{2}$,使尺度滤波器系数之和为1,则范数为1/2。在 Matlab 中,函数 dbaux 和 dbwavf 都可用来计算与 Daubechies 小波相关联的尺度滤波器系数,dbaux 可由用户指定尺度滤波器系数之和,而 dbwavf 则限定尺度滤波器系数之和为1。例如,由 dbaux(2,1)和 dbwavf('db2')都返回与二阶 Daubechies 小波相关联的尺度滤波器,且其系数之和为1。Daubechies 小波和尺度函数都没有解析表达式,只能通过迭代计算画出它们的波形。迭代计算使用多分辨率分析的重构算法,其思路如下:设想多分辨率分析中的信号就是尺度函数,那么在某一分辨率时离散逼近信号中只有一个值不为零,而在同一分辨率及更精细的分辨率时离散细节信号全为零。例如$f(t)=\phi(t)$,则只有$a_0(0)=1$,而其他的$a_0(n)$值全为零,即$a_0(n)=\delta(n)$;所有的$\{d_j(n), j<0\}$全部等于零。通过一级重构,可以算出$a_{-1}(n)$;再通过一级重构计算,得$a_{-2}(n)$;多次迭代计算就得到尺度函数足够密集的抽样,从而画出它的波形。小波函数也可按相似的迭代计算画出。Matlab 中的函数 wavefun 就是通过迭代计算画出小波和尺度函数波形的。图4.1为4阶 Daubechies 小波及尺度函数。

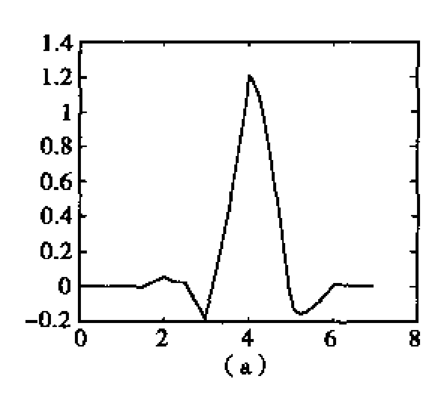
4.4　Symmlets 和 Coiflets 小波

4.4.1　Symmlets 小波

在设计 Daubechies 滤波器时,选取$Q(z)$的所有零点均位于单位圆内,于是$h(n)$的能量主要集中在$n=0$附近,从而$h(n)$很不对称,不具有线性相位,因此$\phi(t)$和$\psi(t)$也很不对称。Daubechies 已经证明[8]:除哈尔滤波器之外,所有实系数紧支尺度滤波器都不对称,不具有线性相位。Daubechies 在设计 Symmlets 滤波器时,优化选择$Q(z)$的零点使$h(n)$几乎具有线性相位,从而得到近似对称或反对称的小波和尺度函数。K阶 Symmlets 小波具有K阶消失矩,支集为$[-K+1,K]$,滤波器长度为$2K$。图4.2画出了4阶 Symmlets 小波及尺度函数。

4.4.2　Coiflets 小波

设计 Daubechies 小波时,只考虑了小波获得最大消失矩,而未考虑尺度函数的消失矩。现



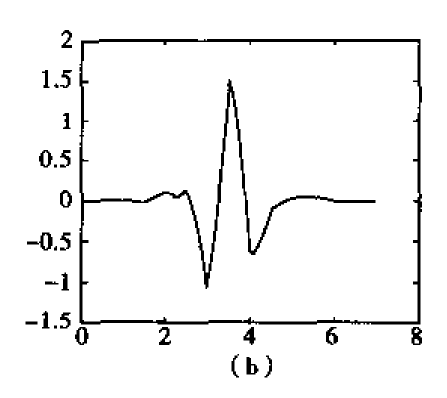


图 4.2 sym4 尺度函数(a)及小波(b)

在同时考虑尺度函数和小波的消失矩，即

$$\int x^k \psi(x)\mathrm{d}x = 0, \quad k = 0,1,\cdots,K-1 \tag{4.4.1}$$

$$\int \phi(x)\mathrm{d}x = 1, \quad \int x^k \phi(x)\mathrm{d}x = 0, \quad k = 1,\cdots,K-1 \tag{4.4.2}$$

对某一精细尺度 2^J，如信号 $f(t)$ 在 $t=2^J n$ 的邻域内是 $C^k, k<K$ 的，则 $f(t)$ 可在此邻域内展开为 k 次泰勒多项式，由(4.4.2)式推知

$$2^{-J/2}\langle f,\phi_{J,n}\rangle \cong f(2^J n) + O(2^{(k+1)J}) \tag{4.4.3}$$

上式意味着信号的采样值是其尺度系数(多分辨率分析中的离散逼近信号)的高阶近似，而且 Coiflets 小波的消失矩阶数越高，用信号的采样值逼近多分辨率分析中的离散逼近信号的误差越小。此外，Coiflets 尺度函数和小波将更为对称。在某些实际应用中，对称性或许更加重要。

根据傅立叶变换的定义，(4.4.2)式等价于

$$\left.\frac{\mathrm{d}^k}{\mathrm{d}\omega^k}\hat{\phi}(\omega)\right|_{\omega=0} = 0, k = 1,\cdots,K-1 \tag{4.4.4}$$

由二尺度方程 $\hat{\phi}(2\omega) = \frac{1}{\sqrt{2}}H(\omega)\hat{\phi}(\omega)$，有

$$2\hat{\phi}'(2\omega) = \frac{1}{\sqrt{2}}[H'(\omega)\phi(\omega) + H(\omega)\phi'(\omega)]$$

由于 $\hat{\phi}'(0)=0$，而 $\phi(0)\neq 0$，故由上式推知 $H'(0)=0$。以此类推，知(4.4.4)式意味着

$$\left.\frac{\mathrm{d}^k}{\mathrm{d}\omega^k}H(\omega)\right|_{\omega=0} = 0, k = 1,\cdots,K-1 \tag{4.4.5}$$

于是 $H(\omega)$ 可写成

$$H(\omega) = \sqrt{2}[1 + (1-\mathrm{e}^{-i\omega})^K Q_2(\omega)] \tag{4.4.6}$$

其中，$Q_2(\omega)$ 是一个三角多项式。显然，上式满足(4.1.9)式。

在构造 Daubechies 小波时已指出(4.4.1)式意味着

$$H(\omega) = \sqrt{2}\left(\frac{1+\mathrm{e}^{-i\omega}}{2}\right)^K Q_1(\omega) \tag{4.4.7}$$

考虑 K 为偶数，即 $K=2L$ 的情况，这时(4.4.6)式和(4.4.7)式可写成

$$H(\omega) = \sqrt{2}[1 + (-1)^L 2^{2L}\mathrm{e}^{-iL\omega}\left(\sin^2(\omega/2)\right)^L Q_2(\omega)]$$

$$H(\omega) = \sqrt{2}\mathrm{e}^{-iL\omega}\left(\cos^2(\omega/2)\right)^L Q_1(\omega)$$

令 $P_2(\omega)=(-1)^L 2^{2L} e^{-iL\omega} Q_2(\omega)$，$P_1(\omega)=e^{-iL\omega}Q_1(\omega)$，则有

$$\left(\cos^2\frac{\omega}{2}\right)^L P_1(\omega) = 1+\left(\sin^2\frac{\omega}{2}\right)^L P_2(\omega) \tag{4.4.8}$$

这时(4.4.6)式和(4.4.7)式意味着必须找到满足(4.4.8)式的两个三角多项式 P_1 和 P_2。根据 Bezout 定理，

$$P_1(\omega) = \sum_{l=0}^{L-1}\binom{L-1+l}{l}\left(\sin^2\frac{\omega}{2}\right)^l + \left(\sin^2\frac{\omega}{2}\right)^L f(\omega) \tag{4.4.9}$$

其中，f 是一任意三角多项式，它的设计必须使尺度滤波器的傅立叶变换满足正交条件 $|H(\omega)|^2+|H(\omega+\pi)|^2=2$。

在 Matlab 中，Coiflets 小波的阶定义为 L，所以 L 阶 Coiflets 小波的消失矩为 $2L$，滤波器的长度为 $6L$。与 Daubechies 小波相比，在同样的小波消失矩时，与 Coiflets 小波相关联的滤波器有更长的长度，但尺度函数和小波都非常接近对称，而且尺度函数具有 $2L-1$ 阶消失矩。可以用函数 coifwavf 计算与 coifL 小波相关联的尺度滤波器；也可以用 orthfilt 函数计算与 coifL 小波相关联的滤波器组；还可以用 wavefun 画出与 coifL 小波相关联的尺度函数和小波。图 4.3 绘制了与 coif4 相关联的尺度函数和小波，可以明显地看到尺度函数和小波都非常接近对称。

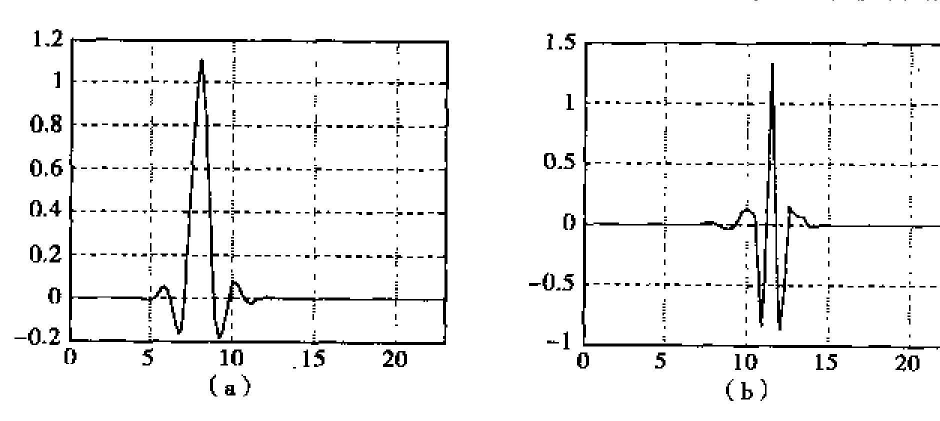


图 4.3 coif4 (a)尺度函数

4.5 样条小波

在多分辨率分析中，曾简要地提到过[0,1)区间的特征函数是多分辨率分析的一个尺度函数，与它对应的是哈尔小波。其实那只是本节要讨论的样条小波的一种最简单的情况。样条函数是一类分段光滑又在各段交接处具有一定光滑性的函数，而在小波分析中用得最多的是基数 B-样条函数，因为它具有最小可能的支撑宽度，又便于算法的计算机处理和实时实现。所以下面只讨论基数 B-样条。

分段常数空间记为 S_0，分段多项式空间记为 S_m，m 为多项式的次数。当 m 为正整数时，S_m 称为基数样条空间，也是讨论样条小波的基本空间。基数样条空间 S_m 是所有 $f\in C^{m-1}(R)$ 的集合，f 在任一区间 $(k,k+1)_{k\in Z}$ 是不超过 m 次的多项式，在整数节点处具有连续的 $m-1$ 阶导数。

零次 B-样条其实就是[0,1)区间的特征函数，记作 $N_0(x)$

$$N_0(x) = \begin{cases} 1, & x \in [0,1) \\ 0, & x \notin [0,1) \end{cases} \tag{4.5.1}$$

m 次基数 B-样条 $N_m(x)$ 由如下卷积迭代关系定义

$$N_m(x) = N_{m-1}(x) * N_0(x) = \int_0^1 N_{m-1}(x-t)\mathrm{d}t \quad m \geq 1 \tag{4.5.2}$$

也就是说，m 次基数 B-样条 $N_m(x)$ 是零次 B-样条 $N_0(x)$ 自身作 m 次卷积的结果。显然，$N_m(x)$ 是紧支的，其支撑宽度为

$$\mathrm{supp}N_m = [0, m+1] \tag{4.5.3}$$

其支撑中心为 $(m+1)/2$，它也是 $N_m(x)$ 的对称中心，即

$$N_m(\frac{m+1}{2} + x) = N_m(\frac{m+1}{2} - x) \tag{4.5.4}$$

且 m 次基数 B-样条 $N_m(x)$ 具有如下性质：

$$N_m(x) > 0, \quad 0 < x < m+1 \tag{4.5.5}$$

$$\sum_{k=-\infty}^{\infty} N_m(x-k) = 1, \quad \forall x \in R \tag{4.5.6}$$

$$N'_m(x) = \Delta N_{m-1}(x) = N_{m-1}(x) - N_{m-1}(x-1) \tag{4.5.7}$$

$$N_m(x) = \frac{x}{m}N_{m-1}(x) + \frac{m+1-x}{m}N_{m-1}(x-1) \tag{4.5.8}$$

下面讨论基数 B-样条在频域的性质。由(4.5.1)式很容易得到零次基数 B-样条的傅立叶变换为

$$\hat{N}_0(\omega) = \mathrm{e}^{-i\omega/2}\frac{\sin(\omega/2)}{(\omega/2)} \tag{4.5.9}$$

从而由卷积迭代关系(4.5.2)式得 m 次基数 B 样条的傅立叶变换为

$$\hat{N}_m(\omega) = \mathrm{e}^{-i\frac{m+1}{2}\omega}\left(\frac{\sin(\omega/2)}{(\omega/2)}\right)^{m+1} \tag{4.5.10}$$

如将 m 次基数 B-样条 $N_m(x)$ 平移为 $N_m[x+(m+1)/2]$，即对称中心移至坐标原点，傅立叶变换将不再出现上式中的线性相位，但 m 为偶数时将出现半整数节点。为避免出现半整数节点，m 为偶数时将 $N_m(x)$ 平移为 $N_m(x+m/2)$，对称中心移至 1/2。将 $N_m(x)$ 整数平移后的样条函数记为 $\theta_m(x)$，则其傅立叶变换为

$$\hat{\theta}_m(\omega) = \mathrm{e}^{-i\frac{\varepsilon\omega}{2}}\left(\frac{\sin(\omega/2)}{(\omega/2)}\right)^{m+1} \tag{4.5.11}$$

上式中，当 m 为奇数时，$\varepsilon=0$，$\theta_m(x)$ 的支集为 $[-(m+1)/2,(m+1)/2]$，对称中心为 0；当 m 为偶数时，$\varepsilon=1$，$\theta_m(x)$ 的支集为 $[-m/2,(m/2)+1]$，对称中心为 1/2。这便是第 3 章所讨论的样条多分辨率逼近。

图 4.4(a)，(b)，(c)，(d) 分别画出了 0 次、1 次、2 次、3 次样条函数。它们的数学表达式为

$$N_0(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{其他} \end{cases}$$

$$N_1(x) = \begin{cases} x, & 0 \leq x < 1 \\ 2-x, & 1 \leq x < 2 \\ 0, & \text{其他} \end{cases}$$

$$N_2(x)=\begin{cases}\frac{1}{2}x^2, & 0\leqslant x<1\\ \frac{3}{4}-\left(x-\frac{3}{2}\right)^2, & 1\leqslant x<2\\ \frac{1}{2}(x-3)^2, & 2\leqslant x<3\\ 0, & \text{其他}\end{cases}$$

$$N_3(x)=\begin{cases}\frac{1}{6}x^3, & 0\leqslant x<1\\ -\frac{1}{2}x^3+2x^2-2x+\frac{2}{3}, & 1\leqslant x<2\\ \frac{1}{2}x^3-4x^2+10x-\frac{22}{3}, & 2\leqslant x<3\\ \frac{1}{6}x^3+2x^2-8x+\frac{32}{3}, & 3\leqslant x<4\\ 0, & \text{其他}\end{cases}$$

可以明显地看到：基数 B-样条在整数区间 $\{[k,k+1]\}_{k\in Z}$ 是不超过 m 次的多项式，在整数节点处 $m-1$ 阶导数连续；次数越高，样条函数越光滑，但其支撑宽度也越宽。

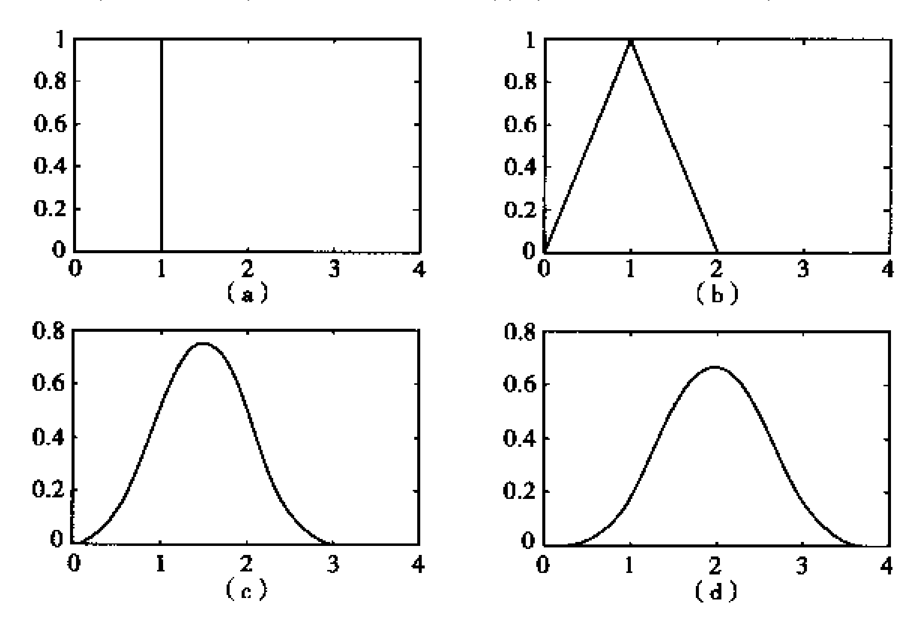


图 4.4　基数 B-样条函数

可以证明[13]：m 次基数 B-样条整数平移形成的函数集 $\{N_m(x-k)\}_{k\in Z}$ 是基数样条空间 S_m 的一个基；而且它还是 V_0 的一个 Riesz 基，所以 $N_m(x)$ 可以作为一个尺度函数而生成一个多分辨率分析。用 $\theta_m(x)$ 作为尺度函数而生成的多分辨率逼近便是第 3 章所讨论的样条多分辨率逼近。很显然，当 $m\geqslant 1$ 时，$\{\theta_m(x-k)\}_{k\in Z}$ 不是一个正交基，但可以用 2.1 节讲过的正交化方法得到正交基。由(4.5.11)式可得

$$\sum_{k=-\infty}^{+\infty}|\hat{\theta}_m(\omega+2k\pi)|^2=(\sin(\omega/2))^{2m+2}\sum_{k=-\infty}^{+\infty}\frac{1}{(\omega/2+k\pi)^{2m+2}}$$

$\theta_m(x)$ 正交化后得到的尺度函数记为 $\phi(x)$，由(2.1.14)式

$$\hat{\phi}(\omega) = \frac{\exp(-i\varepsilon\omega/2)}{\omega^{m+1}\left[\sum_{k=-\infty}^{+\infty}\frac{1}{(\omega+2k\pi)^{2m+2}}\right]^{\frac{1}{2}}} \tag{4.5.12}$$

这正是 Lemarie 和 Battle 提出的多分辨率分析中的尺度函数。

第5章 双正交小波及多分辨率分析

正交小波是一类很重要的小波，尤其是紧支撑的正交小波，如 Daubechies 小波，Symlets 小波和 Coiflets 小波都获得了广泛的应用。但紧支撑的正交小波，除 Haar 小波以外，都不具有对称性。这就意味着与它们相联系的滤波器不是线性相位的 FIR 数字滤波器。在图像处理等诸多应用领域，线性相位滤波器却非常重要。一旦放松正交性，这一矛盾便得到很好的解决，这就是现在要讨论的双正交小波。可以从多分辨率逼近的角度来理解双正交小波，而且可以看到双正交小波和多通道多采样率滤波器组以及子带编码这些技术是有密切联系的。Sweldens 将双正交小波多分辨率逼近加以推广，提出了提升方案，用提升方案构造的小波被称为第二代小波，它们在图像压缩中已得到重要应用。

5.1 双正交小波及多分辨率分析

信号的双正交小波多分辨率逼近是由彼此正交的两个尺度函数 ϕ 和 $\tilde{\phi}$ 生成的两组嵌套空间序列 $\{V_j\}_{j\in Z}$ 和 $\{\tilde{V}\}_{j\in Z}$，两个彼此正交的小波 ψ 和 $\tilde{\psi}$ 二进伸缩和平移构成补空间 W_j 和 $\tilde{W}_j$ 的 Riesz 基。$(\phi,\tilde{\phi},\psi,\tilde{\psi})$ 与双正交滤波器组 $(h,\tilde{h},g,\tilde{g})$ 之间存在着二尺度关系，用双正交滤波器组来实现双正交小波多分辨率分析的 Mallat 算法。

5.1.1 双正交小波多分辨率逼近

双正交小波多分辨率逼近中的两个尺度函数 ϕ 和 $\tilde{\phi}$ 分别生成两个 $L^2(R)$ 的多分辨率逼近 $(V_j)_{j\in Z}$ 和 $(\tilde{V}_j)_{j\in Z}$，但它们平移形成的函数族却不必是 V_0 和 $\tilde{V}_0$ 的正交基，而只要求它们构成 V_0 和 $\tilde{V}_0$ 的 Riesz 基。它们在下述意义下双正交。

定义 5.1 设 ϕ 和 $\tilde{\phi}$ 分别生成两个 $L^2(R)$ 的多分辨率逼近 $(V_j)_{j\in Z}$ 和 $(\tilde{V}_j)_{j\in Z}$，若

$$\langle \phi(t-n),\tilde{\phi}(t-n')\rangle = \delta_{n,n'} \tag{5.1.1}$$

则称 ϕ 和 $\tilde{\phi}$ 是双正交尺度函数。

根据多分辨率逼近的定义，ϕ 和 $\tilde{\phi}$ 生成 $L^2(R)$ 的多分辨率逼近嵌套空间序列

$$\cdots\subset V_2\subset V_1\subset V_0\subset V_{-1}\subset V_{-2}\cdots \tag{5.1.2}$$

$$\cdots\subset \tilde{V}_2\subset \tilde{V}_1\subset \tilde{V}_0\subset \tilde{V}_{-1}\subset \tilde{V}_{-2}\cdots \tag{5.1.3}$$

其中

$$\tilde{V}_j = \overline{\text{span}\{\tilde{\phi}_{j,n}\}} \tag{5.1.4}$$

$$V_j = \overline{\text{span}\{\phi_{j,n}\}} \tag{5.1.5}$$

$\tilde{\phi}$ 平移形成的函数族构成 $\tilde{V}_0$ 的 Riesz 基意味着，对任意 $f(t)\in\tilde{V}_0$，有

$$f(t) = \sum_n a_0(n)\tilde{\phi}_{0,n}(t) \tag{5.1.6}$$

由双正交的定义(5.1.1)式得

$$a_0(n) = \langle f(t),\phi_{0,n}(t)\rangle \tag{5.1.7}$$

实际上，对 $\forall j\in Z$，都可导出类似的关系式，这就是 ϕ 和 $\tilde{\phi}$ 双正交的含义。

设 W_j 是 V_j 在 V_{j-1} 中的补，$\tilde{W}_j$ 是 $\tilde{V}_j$ 在 $\tilde{V}_{j-1}$ 中的补，即

$$V_{j-1} = V_j\oplus W_j, j\in Z \tag{5.1.8}$$

$$\tilde{V}_{j-1} = \tilde{V}_j\oplus \tilde{W}_j, j\in Z \tag{5.1.9}$$

注意，不要求 W_j 与 V_j 正交，$\tilde{W}_j$ 与 $\tilde{V}_j$ 正交。但要求

$$W_j \perp \tilde{V}_j \tag{5.1.10}$$

$$\tilde{W}_j \perp V_j \tag{5.1.11}$$

与正交小波多分辨率逼近类似，嵌套空间序列的因果关系要求尺度函数满足二尺度方程：

$$\phi(t) = \sqrt{2}\sum_k h(k)\phi(2t-k) \tag{5.1.12}$$

$$\tilde{\phi}(t) = \sqrt{2}\sum_k \tilde{h}(k)\tilde{\phi}(2t-k) \tag{5.1.13}$$

$$\psi(t) = \sqrt{2}\sum_k g(k)\phi(2t-k) \tag{5.1.14}$$

$$\tilde{\psi}(t) = \sqrt{2}\sum_k \tilde{g}(k)\tilde{\phi}(2t-k) \tag{5.1.15}$$

$\{\psi_{j,n}\}_{n\in Z}$ 和 $\{\tilde{\psi}_{j,n}\}_{n\in Z}$ 分别构成 W_j 和 $\tilde{W}_j$ 的 Riesz 基，且

$$\langle\psi_{j,n},\tilde{\psi}_{j',n'}\rangle = \delta_{j,j'}\delta_{n,n'} \tag{5.1.16}$$

所以称 ψ 和 $\tilde{\psi}$ 为双正交小波。

对任意的 $f(t)\in L^2(R)$，有

$$f(t) = \sum_{j,n} c_{j,n}\tilde{\psi}_{j,n}(t) \tag{5.1.17}$$

由双正交小波的定义(5.1.16)式可得

$$c_{j,n} = \langle f(t),\psi_{j,n}(t)\rangle \tag{5.1.18}$$

从框架理论的角度看，ψ 和 $\tilde{\psi}$ 互为对偶函数。就是说，也可以用 $\tilde{\psi}$ 进行分解

$$\tilde{c}_{j,n} = \langle f(t), \tilde{\psi}_{j,n}(t) \rangle \tag{5.1.19}$$

而用 ψ 进行重构

$$f(t) = \sum_{j,n} \tilde{c}_{j,n} \psi_{j,n}(t) \tag{5.1.20}$$

从纯数学角度看，(5.1.17)式和(5.1.20)式是等价的。但在实际应用中，例如在图像数据压缩中，要求分解小波具有较高阶的消失矩（或者说消失矩很小），这样系数中有许多很小可以忽略，从而达到高的压缩比，与此同时，重构小波应具有更好的正则性，从而得到较好的重构图像。

5.1.2　双正交滤波器组

与双正交尺度函数和小波 $\{\phi, \tilde{\phi}, \psi, \tilde{\psi}\}$ 关联的4个数字滤波器 $\{h, \tilde{h}, g, \tilde{g}\}$ 称为双正交滤波器组。对二尺度方程(5.1.12)～(5.1.15)两边作傅立叶变换有

$$\hat{\phi}(2\omega) = \frac{1}{\sqrt{2}} H(\omega) \hat{\phi}(\omega) \tag{5.1.21}$$

$$\hat{\tilde{\phi}}(2\omega) = \frac{1}{\sqrt{2}} \tilde{H}(\omega) \hat{\tilde{\phi}}(\omega) \tag{5.1.22}$$

$$\hat{\psi}(2\omega) = \frac{1}{\sqrt{2}} G(\omega) \hat{\phi}(\omega) \tag{5.1.23}$$

$$\hat{\tilde{\psi}}(2\omega) = \frac{1}{\sqrt{2}} \tilde{G}(\omega) \hat{\tilde{\phi}}(\omega) \tag{5.1.24}$$

双正交滤波器组必须满足一定的约束条件。与定理2.1的推导类似，满足双正交条件(5.1.1)的充分必要条件是

$$\sum_{k=-\infty}^{+\infty} \hat{\phi}(\omega + 2k\pi) \hat{\tilde{\phi}}^{*}(\omega + 2k\pi) = 1 \tag{5.1.25}$$

将(5.1.21)式和(5.1.22)式代入上式，并将求和分为对偶数与奇数求和，可以得到(5.1.25)式的等价表达式

$$H^{*}(\omega) \tilde{H}(\omega) + H^{*}(\omega + \pi) \tilde{H}(\omega + \pi) = 2 \tag{5.1.26}$$

5.1.3　双正交小波多分辨率分析的 Mallat 算法

与正交小波多分辨率分析类似，定义

$$a_j(n) = \langle f(t), \phi_{j,n}(t) \rangle \tag{5.1.27}$$

$$d_j(n) = \langle f(t), \psi_{j,n}(t) \rangle \tag{5.1.28}$$

定义 $\bar{h}(n) = h(-n)$，则由二尺度方程(5.1.12)可得

$$\phi_{j+1,n}(t) = \sum_{k} \bar{h}(k) \phi_{j,2n-k}(t)$$

将 $f(t)$ 与上式两边求内积，有

$$a_{j+1}(n) = \sum_{k} \bar{h}(k) a_j(2n-k) \tag{5.1.29}$$

类似的，定义 $\bar{g}(n)=g(-n)$，则由二尺度方程(5.1.14)可得

$$d_{j+1}(n)=\sum_k \bar{g}(k)a_j(2n-k) \tag{5.1.30}$$

(5.1.29)式和(5.1.30)式就是双正交小波多分辨率分析的 Mallat 分解算法。

由上述双正交性知

$$\tilde{f}_{A_j}(t)=\sum_n a_j(n)\tilde{\phi}_{j,n}(t) \tag{5.1.31}$$

属于 $\tilde{V}_j$，可以理解为$f(t)$在 $\tilde{V}_j$ 的 Riesz 基上的线性展开，但绝不是正交投影。类似地，

$$\tilde{f}_{D_j}(t)=\sum_n d_j(n)\tilde{\psi}_{j,n}(t) \tag{5.1.32}$$

属于 $\tilde{W}_j$，可以理解为$f(t)$在 $\tilde{W}_j$ 的 Riesz 基上的线性展开，但也绝不是正交投影。

由于 $\tilde{V}_j=\tilde{V}_{j+1}\oplus\tilde{W}_{j+1}$，所以有

$$\tilde{f}_{A_j}(t)=\tilde{f}_{A_{j+1}}(t)+\tilde{f}_{D_{j+1}}(t) \tag{5.1.33}$$

这意味着

$$\sum_k a_j(k)\tilde{\phi}_{j,k}(t)=\sum_k a_{j+1}(k)\tilde{\phi}_{j+1,k}(t)+\sum_k d_{j+1}(k)\tilde{\psi}_{j+1,k}(t)$$

将 $\phi_{j,n}(t)$与上式两边求内积，有

$$a_j(n)=\sum_k a_{j+1}(k)\langle\phi_{j,n}(t),\tilde{\phi}_{j+1,k}(t)\rangle+\sum_k d_{j+1}(k)\langle\phi_{j,n}(t),\tilde{\psi}_{j+1,k}(t)\rangle \tag{5.1.34}$$

由二尺度方程(5.1.13)和(5.1.15)，有

$$\tilde{\phi}_{j+1,k}(t)=\sum_l \tilde{h}(l-2k)\tilde{\phi}_{j,l}(t)$$

$$\tilde{\psi}_{j+1,k}(t)=\sum_l \tilde{g}(l-2k)\tilde{\phi}_{j,l}(t)$$

将它们代入(5.1.34)式，得

$$a_j(n)=\sum_k \tilde{h}(n-2k)a_{j+1}(k)+\sum_k \tilde{g}(n-2k)d_{j+1}(k) \tag{5.1.35}$$

它就是双正交小波多分辨率分析的 Mallat 重构算法。

综上所述，在连续时间域内用 ϕ 和 ψ 进行分解，用 $\tilde{\phi}$ 和 $\tilde{\psi}$ 进行重构；则在离散时间域内用 h 和 g 进行分解，用 $\tilde{h}$ 和 $\tilde{g}$ 进行重构。

5.2 双通道多采样率滤波器组

上面从多分辨率分析的角度讨论了双正交小波和双正交滤波器组。其实也可以从数字信号处理中的双通道多采样率滤波器组的理想重构条件来理解双正交滤波器组。

图5.1为双通道多采样率滤波器组。其中 $\bar{h}$ 和 $\bar{g}$ 是 h 和 g 的镜像对称，$\bar{h}(n)=h(-n)$，$\bar{g}(n)=g(-n)$。它们分别与输入信号 $a_0(n)$做卷积，然后下采样

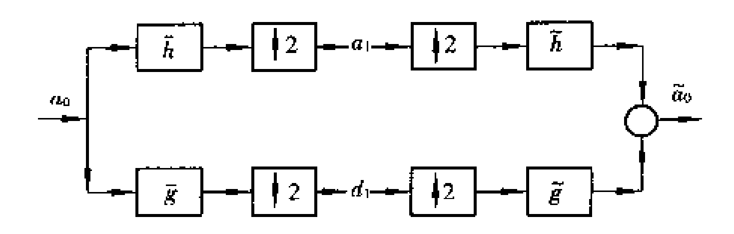


图 5.1　双通道滤波器组

$$a_1(n) = \sum_k h(k-2n)a_0(k) \tag{5.2.1}$$

$$d_1(n) = \sum_k g(k-2n)a_0(k) \tag{5.2.2}$$

其中，$a_1(n)$包含信号 $a_0(n)$的低频信息，而 $d_1(n)$包含信号 $a_0(n)$的高频信息。对 $a_1(n)$和 $d_1(n)$插零之后，分别用对偶低通滤波器 $\tilde{h}(n)$和对偶高通滤波器 $\tilde{g}(n)$滤波，得到重构信号 $\tilde{a}_0(n)$，即

$$\tilde{a}_0(n) = \sum_l [\tilde{h}(n-2l)a_1(l) + \tilde{g}(n-2l)d_1(l)] \tag{5.2.3}$$

对正交小波多分辨率分析的情况，$\tilde{h}=h$，$\tilde{g}=g$ 且 $g(n)=(-1)^{1-n}h(1-n)$，可以实现信号的理想重构，即 $\tilde{a}_0=a_0$。但 FIR 滤波器组不具有线性相位。为了解决线性相位和理想重构之间的矛盾，必须放松对正交性的要求，这就是现在要讨论的双正交小波多分辨率分析。在双正交小波多分辨率分析的滤波器组中，$\tilde{h}$ 和 h 是不同的，但它们却具有线性相位。

当滤波器组满足某些条件时，可实现信号的理想重构，即 $\tilde{a}=a_0$，这些条件称为理想重构条件。实际上，当 $\tilde{a}_0$ 相对于 a_0 有一定延迟，即 $\tilde{a}(n)=\alpha a_0(n-k)$，$k\in Z$ 时，也认为实现了信号的理想重构。

首先在频率域和 z 变换域讨论下采样和插零。设抽取 $x(n)$的偶数样本得到 $x'(n)$，即 $x'(n)=x(2n)$，则有

$$\begin{aligned} X'(z^2) &= \sum_{n\in Z} x'(n)z^{-2n} = \sum_{n\in Z} x(2n)z^{-2n} \\ &= \sum_{k\in Z} \frac{1}{2}[x(k) + (-1)^k x(k)]z^{-k} \\ &= \frac{1}{2}[X(z) + X(-z)] \end{aligned} \tag{5.2.4}$$

令 $z=e^{i\omega}$，由上式可得 $x'(n)$与 $x(n)$的傅立叶变换之间的关系式

$$X'(2\omega) = \frac{1}{2}[X(\omega) + X(\omega+\pi)] \tag{5.2.5}$$

其中第二项产生频谱折叠，重构时必须将这个“混叠”消除掉。

设 $y(n)$是对 $x'(n)$插零的结果，即

$$y(n) = \begin{cases} x'(k), & n = 2k \\ 0, & n = 2k+1 \end{cases}$$

则有

$$Y(z) = \sum_{n\in Z} y(n)z^{-n} = \sum_{k\in Z} x'(k)(z^2)^{-k} = X'(z^2) \tag{5.2.6}$$

$$Y(\omega) = X'(2\omega) \tag{5.2.7}$$

如对 $x(n)$ 先下采样得到 $x'(n)$，再对 $x'(n)$ 插零得到 $y(n)$，则有

$$Y(z) = X'(z^2) = \frac{1}{2}[X(z) + X(-z)] \tag{5.2.8}$$

$$Y(\omega) = X'(2\omega) = \frac{1}{2}[X(\omega) + X(\omega + \pi)] \tag{5.2.9}$$

下述定理给出保证理想重构的充分必要条件。

定理 5.1 滤波器组对任何输入信号实现理想重构当且仅当

$$H^*(\omega + \pi)\tilde{H}(\omega) + G^*(\omega + \pi)\tilde{G}(\omega) = 0 \tag{5.2.10}$$

和

$$H^*(\omega)\tilde{H}(\omega) + G^*(\omega)\tilde{G}(\omega) = 2 \tag{5.2.11}$$

证明 由图 5.1 及下采样和插零的 z 变换表达式(5.2.8)，有

$$\begin{aligned}\tilde{A}_0(z) &= \frac{1}{2}[\overline{H}(z)A_0(z) + \overline{H}(-z)A_0(-z)]\tilde{H}(z) + \\ &\quad \frac{1}{2}[\overline{G}(z)A_0(z) + \overline{G}(-z)A_0(-z)]\tilde{G}(z) \\ &= \frac{1}{2}[\overline{H}(z)\tilde{H}(z) + \overline{G}(z)\tilde{G}(z)]A_0(z) + \\ &\quad \frac{1}{2}[\overline{H}(-z)\tilde{H}(z) + \overline{G}(-z)\tilde{G}(z)]A_0(-z)\end{aligned} \tag{5.2.12}$$

在时域用上横线表示镜像对称，故在 z 变换域意味着取 z 的倒数，即 $\overline{H}(z) = H(z^{-1})$；进一步，如 h 为实系数，则意味着傅立叶变换取共轭，即 $\overline{H}(\omega) = H^*(\omega)$。(5.2.12)式第二项对应着因下取样而造成的混叠，为实现理想重构，这一项应等于零。于是，z 变换的理想重构条件表述为

$$\overline{H}(-z)\tilde{H}(z) + \overline{G}(-z)\tilde{G}(z) = 0 \tag{5.2.13}$$

$$\overline{H}(z)\tilde{H}(z) + \overline{G}(z)\tilde{G}(z) = 2 \tag{5.2.14}$$

从而，由上列两式可得(5.2.10)式和(5.2.11)式。证毕。

由定理 5.1 知重构滤波器 $\tilde{h}$ 和 $\tilde{g}$ 完全被分解滤波器 h 和 g 决定。可将理想重构条件(5.2.10)式和(5.2.11)式写成如下矩阵形式：

$$\begin{pmatrix} H(\omega) & G(\omega) \\ H(\omega + \pi) & G(\omega + \pi) \end{pmatrix}\begin{pmatrix} \tilde{H}^*(\omega) \\ \tilde{G}^*(\omega) \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \tag{5.2.15}$$

从而得

$$\begin{pmatrix} \tilde{H}^*(\omega) \\ \tilde{G}^*(\omega) \end{pmatrix} = \frac{2}{\Delta(\omega)}\begin{pmatrix} G(\omega + \pi) \\ -H(\omega + \pi) \end{pmatrix} \tag{5.2.16}$$

其中

$$\Delta(\omega) = H(\omega)G(\omega + \pi) - H(\omega + \pi)G(\omega) \tag{5.2.17}$$

上列两式更清楚地表示出分解滤波器和重构滤波器之间的关系。仅当 $\omega \in [-\pi, +\pi]$ 时，$\Delta(\omega) \neq 0$，重构滤波器才是稳定的。对 FIR 滤波器组，$\Delta(\omega)$ 是可求值的。

对 FIR 滤波器组，有下述定理。

定理 5.2　理想重构滤波器满足

$$H^*(\omega)\tilde{H}(\omega)+H^*(\omega+\pi)\tilde{H}(\omega+\pi)=2 \tag{5.2.18}$$

如果滤波器为 FIR 滤波器，那么，存在 $a\in R,l\in Z$，使得

$$G(\omega)=a\mathrm{e}^{-i(2l+1)\omega}\tilde{H}^*(\omega+\pi)$$

$$\tilde{G}(\omega)=a^{-1}\mathrm{e}^{-i(2l+1)\omega}H^*(\omega+\pi) \tag{5.2.19}$$

证明　由于采用 FIR 滤波器，所以 $\overline{H}$ 和 $\overline{G}$ 都是 z 的多项式。(5.2.14)式不容许 $\overline{H}$ 和 $\overline{G}$ 有共同的零点；于是，(5.2.13)式意味着

$$\tilde{G}(z)=\overline{H}(-z)P(z),\qquad \tilde{H}(z)=-\overline{G}(-z)P(z) \tag{5.2.20}$$

其中，$P(z)$ 是一个 z 的多项式。将(5.2.20)式代回到(5.2.14)式得

$$[H(-z)G(z)-H(z)G(-z)]\overline{P}(z)=2$$

由于能除常数的惟一多项式必然是单项式，故

$$P(z)=a^{-1}z^k,a\in R,k\in Z$$

从而(5.2.20)式可写成

$$\tilde{G}(z)=a^{-1}z^k\overline{H}(-z),\qquad G(z)=-a(-1)^kz^k\bar{\tilde{H}}(-z)$$

为使 G 和 $\tilde{G}$ 有对称性，可取 $k=-(2l+1)$，于是

$$\tilde{G}(z)=a^{-1}z^{-(2l+1)}\overline{H}(-z),\qquad G(z)=az^{-(2l+1)}\bar{\tilde{H}}(-z) \tag{5.2.21}$$

再把(5.2.21)式代入(5.2.14)式得

$$\overline{H}(z)\tilde{H}(z)+\overline{H}(-z)\tilde{H}(-z)=2 \tag{5.2.22}$$

(5.2.22)式和(5.2.21)式就是满足理想重构条件的滤波器组在 z 变换域的表达式。当滤波器为实系数时，由它们便可得到理想重构条件的频域表达式(5.2.18)式和(5.2.19)式。证毕。

因子 a 是放大率，分解滤波器和重构滤波器的放大率互为倒数。l 表示相位延迟。通常取 $a=1,l=0$。这时，有

$$G(\omega)=\mathrm{e}^{-i\omega}\tilde{H}^*(\omega+\pi)$$

$$\tilde{G}(\omega)=\mathrm{e}^{-i\omega}H^*(\omega+\pi) \tag{5.2.23}$$

由(5.2.18)式和(5.2.23)式可得时间域的理想重构条件为

$$\sum_n \tilde{h}(n)h(n-2k)=\delta(k) \tag{5.2.24}$$

和

$$g(n)=(-1)^{1-n}\tilde{h}(1-n) \tag{5.2.25}$$

$$\tilde{g}(n)=(-1)^{1-n}h(1-n) \tag{5.2.26}$$

这时滤波器 (h,g) 和 $(\tilde{h},\tilde{g})$ 是完全对称的，所以也可以用 $(\tilde{h},\tilde{g})$ 作信号分解，用 (h,g) 作信号重构。

当分解滤波器等于重构滤波器，即 $\tilde{h}=h$ 时，(5.2.24)式退化为

$$\sum_n h(n)h(n-2k) = \delta(k)$$

这正是正交小波多分辨率分析中尺度滤波器 h 的正交性表达式。图 5.1 双通道滤波器组成为正交小波多分辨率分析的分解和重构。所以正交实际上是双正交的特殊情况，或者说双正交是正交的推广。

下述定理指出分解滤波器(h,g)和重构滤波器$(\tilde{h},\tilde{g})$的双正交性。

定理 5.3 如果$(h,\tilde{h},g,\tilde{g})$是理想重构滤波器组，那么，它们在如下意义下是双正交的：

$$\sum_n \tilde{h}(n)h(n-2k) = \delta(k) \tag{5.2.27}$$

$$\sum_n \tilde{g}(n)g(n-2k) = \delta(k) \tag{5.2.28}$$

和

$$\sum_n \tilde{h}(n)g(n-2k) = 0 \tag{5.2.29}$$

$$\sum_n \tilde{g}(n)h(n-2k) = 0 \tag{5.2.30}$$

证明 对于理想重构滤波器，由(5.2.18)式，有

$$\frac{1}{2}[H^*(\omega)\tilde{H}(\omega)+H^*(\omega+\pi)\tilde{H}(\omega+\pi)]=1$$

从下采样的频域表达式(5.2.5)可以看出，上式左边是 $\bar{h}$ 与 $\tilde{h}$ 卷积所得序列下采样的傅立叶变换，所以它在时域的表达式便是(5.2.27)式。

由于 $H(\omega)$及 $\tilde{H}(\omega)$是 2π 的周期函数，由(5.2.23)有

$$G(\omega+\pi) = -e^{-i\omega}\tilde{H}^*(\omega)$$

$$\tilde{G}(\omega+\pi) = -e^{-i\omega}H^*(\omega) \tag{5.2.31}$$

于是可得

$$\frac{1}{2}[G^*(\omega)\tilde{G}(\omega)+G^*(\omega+\pi)\tilde{G}(\omega+\pi)] = 1$$

类似地，它在时域的表达式便是(5.2.28)式。

将(5.2.31)式和(5.2.23)式交叉相乘后有

$$\frac{1}{2}[G^*(\omega)\tilde{H}(\omega)+G^*(\omega+\pi)\tilde{H}(\omega+\pi)] = 0$$

$$\frac{1}{2}[H^*(\omega)\tilde{G}(\omega)+H^*(\omega+\pi)\tilde{G}(\omega+\pi)] = 0$$

同理，它们在时域的表达式便是(5.2.29)式和(5.2.30)式。证毕。

5.3 双正交小波基

从正交尺度滤波器可以构造正交小波，类似地，也可以从双正交滤波器组来构造双正交小波。本节讨论如何从有限冲击响应双正交滤波器组构造紧支双正交小波。

由滤波器(h,g)和$(\tilde{h},\tilde{g})$，可以定义两个尺度函数和小波，它们的傅立叶变换满足

$$\hat{\phi}(2\omega)=\frac{1}{\sqrt{2}}H(\omega)\hat{\phi}(\omega),\hat{\tilde{\phi}}(2\omega)=\frac{1}{\sqrt{2}}\tilde{H}(\omega)\hat{\tilde{\phi}}(\omega) \tag{5.3.1}$$

$$\hat{\psi}(2\omega)=\frac{1}{\sqrt{2}}G(\omega)\hat{\phi}(\omega),\hat{\tilde{\psi}}(2\omega)=\frac{1}{\sqrt{2}}\tilde{G}(\omega)\hat{\tilde{\phi}}(\omega) \tag{5.3.2}$$

在时域，这些二尺度方程式成为

$$\phi(t)=\sqrt{2}\sum_{n}h(n)\phi(2t-n),\tilde{\phi}(t)=\sqrt{2}\sum_{n}\tilde{h}(n)\tilde{\phi}(2t-n) \tag{5.3.3}$$

$$\psi(t)=\sqrt{2}\sum_{n}g(n)\phi(2t-n),\tilde{\psi}(t)=\sqrt{2}\sum_{n}\tilde{g}(n)\tilde{\phi}(2t-n) \tag{5.3.4}$$

与正交情况一样，要求尺度函数的积分等于1，小波的积分等于零，即$\hat{\phi}(0)=\hat{\tilde{\phi}}(0)=1$，$\hat{\psi}(0)=\hat{\tilde{\psi}}(0)=0$。于是，$H(0)=\tilde{H}(0)=\sqrt{2}$，$G(0)=\tilde{G}(0)=0$；从而$H(\pi)=\tilde{H}(\pi)=0$，$|G(\pi)|=|\tilde{G}(\pi)|=\sqrt{2}$。

(5.3.1)～(5.3.4)式意味着，可以由滤波器组生成一对尺度函数和小波。如果h和$\tilde{h}$是冲击响应有限的滤波器，那么，尺度函数和小波是紧支的。尺度函数的傅立叶变换可以写成无限级联的形式

$$\hat{\phi}(\omega)=\prod_{p=1}^{+\infty}\frac{H(2^{-p}\omega)}{\sqrt{2}},\hat{\tilde{\phi}}(\omega)=\prod_{p=1}^{+\infty}\frac{\tilde{H}(2^{-p}\omega)}{\sqrt{2}} \tag{5.3.5}$$

它们是紧支集分布的傅立叶变换，但这些分布可能不会收敛到比较正则的函数，而且它们的能量也可能是无限的。所以必须对滤波器组施加进一步的约束条件以保证$\hat{\phi}$和$\hat{\tilde{\phi}}$是能量有限函数的傅立叶变换。下述定理给出由理想重构滤波器组生成$L^2(R)$的双正交小波基的充分条件。

定理5.4　假设存在严格正的三角多项式$P(e^{i\omega})$和$\tilde{P}(e^{i\omega})$，使得

$$\left|H\left(\frac{\omega}{2}\right)\right|^2P(e^{i\omega/2})+\left|H\left(\frac{\omega}{2}+\pi\right)\right|^2P(e^{i(\omega/2+\pi)})=2P(e^{i\omega}) \tag{5.3.6}$$

$$\left|\tilde{H}\left(\frac{\omega}{2}\right)\right|^2\tilde{P}(e^{i\omega/2})+\left|\tilde{H}\left(\frac{\omega}{2}+\pi\right)\right|^2\tilde{P}(e^{i(\omega/2+\pi)})=2\tilde{P}(e^{i\omega}) \tag{5.3.7}$$

并假定P和$\tilde{P}$是惟一的(规范化后)。设

$$\inf_{\omega\in[-\pi/2,\pi/2]}|H(\omega)|>0,\inf_{\omega\in[-\pi/2,\pi/2]}|\tilde{H}(\omega)|>0 \tag{5.3.8}$$

那么，(5.3.5)式定义的ϕ和$\tilde{\phi}$属于$L^2(R)$，且满足双正交关系

$$\langle\phi(t),\tilde{\phi}(t-n)\rangle=\delta_n \tag{5.3.9}$$

$\{\psi_{j,n}\}_{j,n\in Z}$和$\{\tilde{\psi}_{j,n}\}_{j,n\in Z}$是$L^2(R)$的双正交Riesz基。

小波ψ和$\tilde{\psi}$的双正交性意味着

$$\langle\psi_{j,n}(t),\tilde{\psi}_{l,k}(t)\rangle=\delta_{j,l}\delta_{n,k} \tag{5.3.10}$$

对于双正交Riesz基，任意$f(t)\in L^2(R)$均可分解为

$$f(t) = \sum_{j,n} \langle f(t), \psi_{j,n}(t) \rangle \tilde{\psi}_{j,n}(t) \tag{5.3.11}$$

或

$$f(t) = \sum_{j,n} \langle f(t), \tilde{\psi}_{j,n}(t) \rangle \psi_{j,n}(t) \tag{5.3.12}$$

由 Riesz 稳定性知,存在 $0 < A \leqslant B < +\infty$,使得

$$A\|f\|^2 \leqslant \sum_{j,n} \left| \langle f, \psi_{j,n} \rangle \right|^2 \leqslant B\|f\|^2 \tag{5.3.13}$$

$$\frac{1}{B}\|f\|^2 \leqslant \sum_{j,n} \left| \langle f, \tilde{\psi}_{j,n} \rangle \right|^2 \leqslant \frac{1}{A}\|f\|^2 \tag{5.3.14}$$

由 ϕ 和 $\tilde{\phi}$ 可以生成多分辨率逼近,$\{\phi(t-n)\}_{n\in Z}$是它生成的空间 V_0 的 Riesz 基,而 $\{\tilde{\phi}(t-n)\}_{n\in Z}$是另一空间 $\tilde{V}_0$ 的 Riesz 基。令 V_j 和 $\tilde{V}_j$ 是由二尺度方程

$$g(t) \in V_0 \Leftrightarrow g(2^{-j}t) \in V_j \tag{5.3.15}$$

$$g(t) \in \tilde{V}_0 \Leftrightarrow g(2^{-j}t) \in \tilde{V}_j \tag{5.3.16}$$

定义的空间。可以证明,$\{V_j\}_{j\in Z}$和$\{\tilde{V}_j\}_{j\in Z}$是$L^2(R)$的两个多分辨率逼近。由上述二尺度方程不难看出,$\{\phi_{j,n}\}_{n\in Z}$和$\{\tilde{\phi}_{j,n}\}_{n\in Z}$是 V_j 和 $\tilde{V}_j$ 的 Riesz 基,而$\{\psi_{j,n}\}_{n\in Z}$和$\{\tilde{\psi}_{j,n}\}_{n\in Z}$是两个细节空间 W_j 和 $\tilde{W}_j$ 的 Riesz 基,且

$$V_j = V_{j+1} \oplus W_{j+1} \tag{5.3.17}$$

$$\tilde{V}_j = \tilde{V}_{j+1} \oplus \tilde{W}_{j+1} \tag{5.3.18}$$

分解和重构滤波器的双正交性意味着 $\tilde{W}_j$ 和 V_j 正交,W_j 和 $\tilde{V}_j$ 正交。

5.4 紧支双正交小波

在设计尺度滤波器时,要把正交小波的支集、消失矩和正则性与尺度滤波器的性质联系起来。在设计双正交滤波器时,除了要考虑双正交小波的支集、消失矩和正则性以外,还要考虑双正交小波的对称性。为了构造具有对称或反对称性的紧支双正交小波,双正交滤波器必须是线性相位 FIR 滤波器。

5.4.1 双正交小波的设计

由二尺度方程可以证明:如 h 和 $\tilde{h}$ 的支集是$[N_1, N_2]$和$[\tilde{N}_1, \tilde{N}_2]$,那么,$\phi$ 和 $\tilde{\phi}$ 也有相同的紧支集;而 ψ 和 $\tilde{\psi}$ 的支集则分别为

$$\left[\frac{N_1 - \tilde{N}_2 + 1}{2}, \frac{\tilde{N}_2 - N_1 + 1}{2}\right] \text{和} \left[\frac{\tilde{N}_1 - N_2 + 1}{2}, \frac{N_2 - \tilde{N}_1 + 1}{2}\right] \tag{5.4.1}$$

所以,两个小波的支集长度相同,均为$(N_2 - N_1 + \tilde{N}_2 - \tilde{N}_1)/2$。

ψ 和 $\tilde{\psi}$ 的消失矩阶数取决于 $\tilde{H}(\omega)$和$H(\omega)$在 $\omega = \pi$ 处的零点重数,也即 $\tilde{H}(z)$和 $H(z)$

在 $z=-1$ 处的零点重数。由二尺度方程(5.3.2),如 ψ 有 $\tilde{K}$ 阶消失矩,则 $G(\omega)$ 在 $\omega=0$ 处有 $\tilde{K}$ 重零点。因为 $G(\omega)=e^{-i\omega}\tilde{H}^*(\omega+\pi)$,这意味着 $\tilde{H}(\omega)$ 在 $\omega=\pi$ 处有 $\tilde{K}$ 重零点。类似的,$\tilde{\psi}$ 的消失矩阶数 K 等于 $H(\omega)$ 在 $\omega=\pi$ 处的零点重数。

如 $\tilde{H}(\omega)$ 在 $\omega=\pi$ 处有 $\tilde{K}$ 重零点,则它可以写成如下因式分解的形式

$$\tilde{H}(\omega)=\sqrt{2}\left(\frac{1+e^{-i\omega}}{2}\right)^{\tilde{K}}\tilde{Q}(\omega) \tag{5.4.2}$$

令 $\tilde{B}=\sup_{\omega\in[-\pi,\pi]}|\tilde{Q}(\omega)|$,由定理4.4知 $\tilde{\psi}$ 和 $\tilde{\phi}$ 的一致 Lipschitz $\tilde{\alpha}$ 满足

$$\tilde{\alpha}<\tilde{\alpha}_0=\tilde{K}-\log_2\tilde{B}-1 \tag{5.4.3}$$

一般来说,$\tilde{B}$ 的增加要比 $\tilde{K}$ 的增加慢得多,所以 ψ 的消失矩 $\tilde{K}$ 的增加意味着 $\tilde{\psi}$ 有更好的正则性。类似的,$\tilde{\psi}$ 的消失矩 K 的增加意味着 ψ 有更好的正则性。这正好符合小波分析在数据和图像压缩应用中的要求。

除哈尔小波以外,紧支正交小波没有对称或反对称性。但却可以由线性相位 FIR 滤波器 h 和 $\tilde{h}$ 来生成对称或反对称的紧支双正交小波。

5.4.2　线性相位双正交滤波器

现在讨论一种最简单的情况,滤波器 h 和 $\tilde{h}$ 同时有奇数或偶数个实系数。当它们均有奇数个实系数时,以0为对称中心,很容易证明 ϕ 和 $\tilde{\phi}$ 以0为中心对称,ψ 和 $\tilde{\psi}$ 相对于一个偏移0的中心对称;而当它们均有偶数个实系数时,以1/2为对称中心,ϕ 和 $\tilde{\phi}$ 也以1/2为中心对称,而 ψ 和 $\tilde{\psi}$ 则相对于一个偏移的中心反对称。

当 h 有奇数个实系数,且以0为对称中心时,$h(n)=\bar{h}(n)=h(-n)$,由 z 变换的定义知,$H(z)=\overline{H}(z)=H(z^{-1})$。如果 z_0 是 $H(z)$ 的零点,那么,z_0^{-1} 也必然是 $H(z)$ 的零点;当 h 为实系数时,z_0^* 和 $(z_0^*)^{-1}$ 也是 $H(z)$ 的零点;也就是说,$H(z)$ 的零点是4个成组的。由于这时 $H(z)$ 有偶数个零点,所以如它在 $z=\pm1$ 处有零点,则必然是偶重零点。

当 h 有奇数个实系数,且以0为对称中心时,其傅立叶变换可写成

$$H(\omega)=\sqrt{2}[\cos(\omega/2)]^{2l}Q(\cos\omega) \tag{5.4.4}$$

类似的,当 $\tilde{h}$ 有奇数个实系数,且以0为对称中心时,其傅立叶变换可写成

$$\tilde{H}(\omega)=\sqrt{2}[\cos\omega/2]^{2\tilde{l}}\tilde{Q}(\cos\omega) \tag{5.4.5}$$

将上列两式代入(5.2.18)式,则有

$$[\cos(\omega/2)]^{2L}Q^*(\cos\omega)\tilde{Q}(\cos\omega)+[\sin(\omega/2)]^{2L}Q^*[\cos(\omega+\pi)]\tilde{Q}[\cos(\omega+\pi)]=1 \tag{5.4.6}$$

其中,$L=l+\tilde{l}$。

令 $y=\sin^2(\omega/2)$,且设

$$Q^*(\cos\omega)\tilde{Q}(\cos\omega) = P[\sin^2(\omega/2)] = P(y) \tag{5.4.7}$$

则由(5.4.6)式可得

$$(1-y)^L P(y) + y^L P(1-y) = 1 \tag{5.4.8}$$

当 h 有偶数个实系数,且以 1/2 为对称中心时,$h(n)=h(1-n)=\bar{h}(n-1)$,这时,$H(z)=z^{-1}\overline{H}(z)=z^{-1}H(z^{-1})$。所以 $H(z)$的零点仍然是 4 个成组的。但此时 $H(z)$有奇数个零点,所以如它在 $z=\pm1$ 处有零点,则必然是奇重零点。于是,当 h 和 $\tilde{h}$ 均有偶数个实系数时,它们的傅立叶变换可写成

$$H(\omega) = \sqrt{2}e^{-i\omega/2}[\cos(\omega/2)]^{2l+1}Q(\cos\omega) \tag{5.4.9}$$

和

$$\tilde{H}(\omega) = \sqrt{2}e^{-i\omega/2}[\cos(\omega/2)]^{2\tilde{l}+1}\tilde{Q}(\cos\omega) \tag{5.4.10}$$

将上列两式代入(5.2.18)式,则有

$$[\cos(\omega/2)]^{2L}Q^*(\cos\omega)\tilde{Q}(\cos\omega) + [\sin(\omega/2)]^{2L}Q^*[\cos(\omega+\pi)]\tilde{Q}[\cos(\omega+\pi)] = 1 \tag{5.4.11}$$

其中,$L=l+\tilde{l}+1$。

(5.4.11)式和(5.4.6)式形式完全一样,仅仅是 L 不同而已。仍令 $y=\sin^2(\omega/2)$,则可由(5.4.11)式得到(5.4.8)式。

当 h 和 $\tilde{h}$ 的长度为奇数时,由(5.4.4)式和(5.4.5)式知,$H(\omega)$和 $\tilde{H}(\omega)$在 $\omega=\pi$ 分别有 $K=2l$ 和 $\tilde{K}=2\tilde{l}$ 重零点,于是 $L=l+\tilde{l}=(K+\tilde{K})/2$;而当 h 和 $\tilde{h}$ 的长度为偶数时,由(5.4.9)式和(5.4.10)式知,$H(\omega)$和 $\tilde{H}(\omega)$在 $\omega=\pi$ 分别有 $K=2l+1$ 和 $\tilde{K}=2\tilde{l}+1$ 重零点,于是 $L=l+\tilde{l}+1=(K+\tilde{K})/2$。这样,用 $\tilde{\psi}$ 的消失矩阶数 K 和 ψ 的消失矩阶数 $\tilde{K}$ 时,无论滤波器的长度为奇数或偶数,都有 $L=(K+\tilde{K})/2$。

方程式(5.4.8)和构造 Daubechies 小波时的方程式(4.3.7)形式一样,只不过多项式 P 的定义不同而已。仍然可以用 Bezout 定理得到方程式(5.4.8)的解,然后按(5.4.7)式选择 Q 和 $\tilde{Q}$ 的零点,从而设计出 h 和 $\tilde{h}$。

5.4.3 样条双正交小波

如选择 $\tilde{Q}(\cos\omega)\equiv1$,则由(5.4.7)式知 $Q^*(\cos\omega)=P[\sin^2(\omega/2)]$。代入(5.4.5)式和(5.4.10)式可得

$$\tilde{H}(\omega) = \sqrt{2}e^{-i\varepsilon\omega/2}[\cos(\omega/2)]^{\tilde{K}} \tag{5.4.12}$$

当 $\tilde{K}$ 取偶数时,$\varepsilon=0$;而当它取奇数时,$\varepsilon=1$。与构造 Daubechies 小波类似,选择多项式 $P(y)$ 的最低次解,这样滤波器 h 的长度最短。由(5.4.4)式和(5.4.9)式知其傅立叶变换为

$$H(\omega) = \sqrt{2}e^{-i\varepsilon\omega/2}[\cos(\omega/2)]^K \sum_{q=0}^{L-1}\binom{L-1+q}{q}[\sin(\omega/2)]^{2q} \tag{5.4.13}$$

同样,当 K 取偶数时,$\varepsilon=0$;而当它取奇数时,$\varepsilon=1$。

由二尺度方程(5.3.1)可得

$$\hat{\tilde{\phi}}(\omega) = e^{-\frac{i\varepsilon\omega}{2}}\left(\frac{\sin(\omega/2)}{(\omega/2)}\right)^{\tilde{K}} \tag{5.4.14}$$

它正是 $\tilde{K}-1$ 阶样条函数的傅立叶变换。因为 $\tilde{\psi}$ 是 $\tilde{\phi}(2t-n)$ 的线性组合，所以它是与 $\tilde{\phi}$ 次数相同的紧支集多项式样条。这样生成的小波 $\tilde{\psi}$ 和 ψ 称为双正交样条小波。

从(5.4.12)式和(5.4.13)式可以看到滤波器 h 和 $\tilde{h}$ 的长度分别为 $2K+\tilde{K}-1$ 和 $\tilde{K}+1$，所以 ϕ 和 $\tilde{\phi}$ 的支集宽度分别是 $2K+\tilde{K}-2$ 和 $\tilde{K}$。于是双正交样条小波的支集宽度均为 $K+\tilde{K}-1$。

例 5.1　取 $\tilde{K}=2, K=2$，则 $L=2, \varepsilon=0$。

由(5.4.12)式有

$$\tilde{H}(\omega) = \sqrt{2}[\cos(\omega/2)]^2 = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{4}e^{i\omega} + \frac{\sqrt{2}}{4}e^{-i\omega}$$

于是，$\tilde{h}(0)=\sqrt{2}/2, \tilde{h}(-1)=\tilde{h}(1)=\sqrt{2}/4$。

由(5.4.13)式有

$$\begin{aligned} H(\omega) &= \sqrt{2}[\cos(\omega/2)]^2[1+2\sin^2(\omega/2)] \\ &= \frac{3\sqrt{2}}{4} + \frac{\sqrt{2}}{4}e^{i\omega} + \frac{\sqrt{2}}{4}e^{-i\omega} - \frac{\sqrt{2}}{8}e^{i2\omega} - \frac{\sqrt{2}}{8}e^{-i2\omega} \end{aligned}$$

于是，$h(0)=3\sqrt{2}/4, h(-1)=h(1)=\sqrt{2}/4, h(-2)=h(2)=-\sqrt{2}/8$。

例 5.2　取 $\tilde{K}=3, K=1$，则 $L=2, \varepsilon=1$。

由(5.4.12)式有

$$\begin{aligned} \tilde{H}(\omega) &= \sqrt{2}e^{-i\omega/2}[\cos(\omega/2)]^3 \\ &= \frac{\sqrt{2}}{8}(3+3e^{-i\omega}+e^{i\omega}+e^{-i2\omega}) \end{aligned}$$

于是，$\tilde{h}(0)=\tilde{h}(1)=3\sqrt{2}/8, \tilde{h}(-1)=\tilde{h}(2)=\sqrt{2}/8$。

由(5.4.13)式有

$$\begin{aligned} H(\omega) &= \sqrt{2}e^{-i\omega/2}\cos(\omega/2)[1+2\sin^2(\omega/2)] \\ &= \frac{3\sqrt{2}}{4}(1+e^{-i\omega}) - \frac{\sqrt{2}}{4}(e^{i\omega}+e^{-i2\omega}) \end{aligned}$$

于是，$h(0)=h(1)=3\sqrt{2}/4, h(-1)=h(2)=-\sqrt{2}/4$。

Matlab 中的双正交滤波器组正是按这种方法设计的，这时分解和重构小波的消失矩阶数要么同时为奇数，要么同时为偶数，例如‘bior3.5’重构小波有 3 阶消失矩而分解小波有 5 阶消失矩。用 biorfilt 可以获得双正交滤波器组，图 5.2 为‘bior3.5’双正交滤波器组，其中，(a)和(b)为分解低通和高通，(c)和(d)为重构低通和高通。图 5.3 为‘bior3.5’尺度函数和小波，其中(a)和(b)分别为分解尺度函数和小波，(c)和(d)分别为重构尺度函数和小波。可以明显地看到，分解小波具有较高阶的消失矩，因而振荡次数较多，具有更好的频域局部性；重构小波具有更好的正则性，因而更光滑。

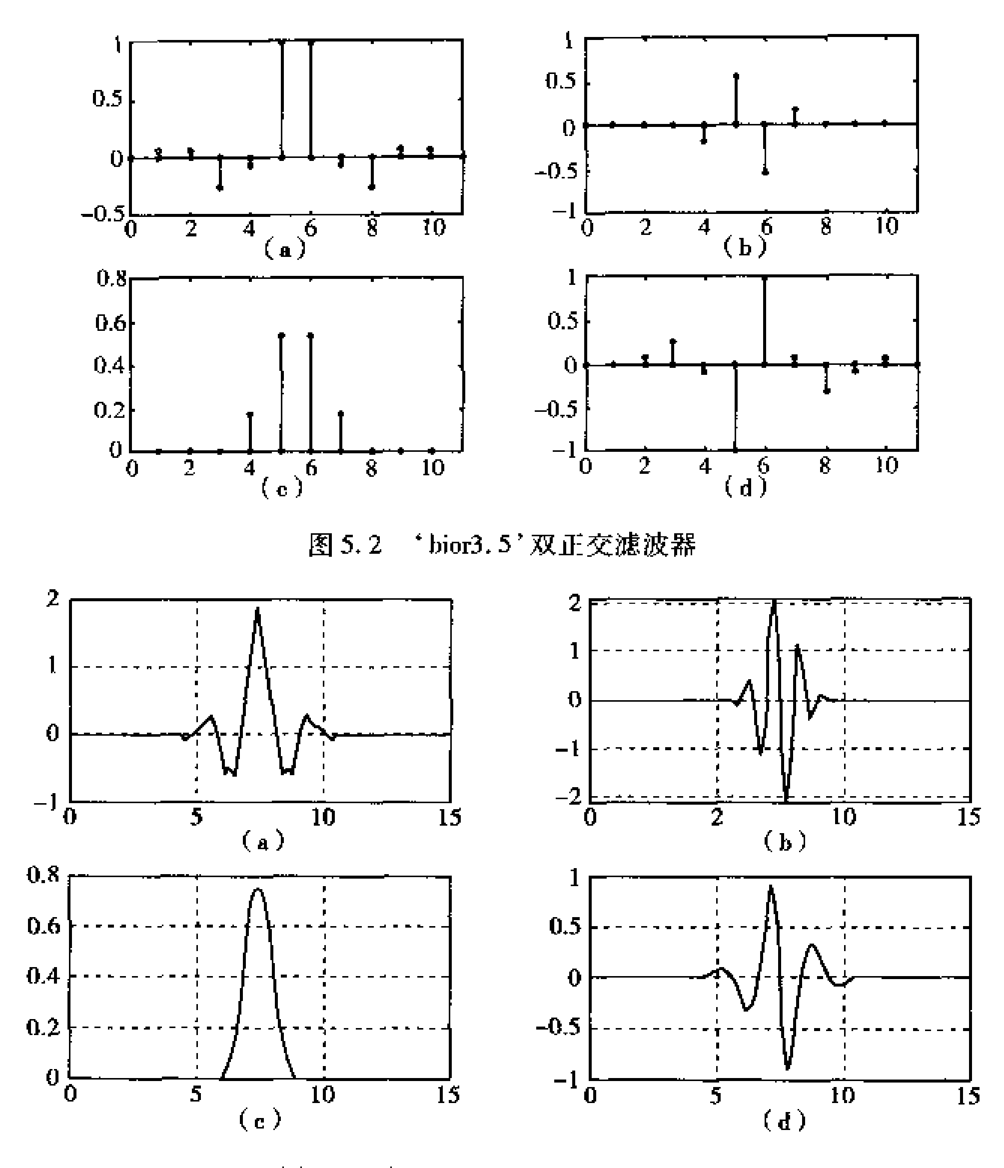


图 5.2 ‘bior3.5’双正交滤波器

图 5.3 ‘bior3.5’双正交小波及尺度

按照上述方法设计的样条双正交小波在获得最大消失矩的同时有最小紧支集，但 h 和 $\tilde{h}$ 有不同的长度。为了得到长度近似相等的滤波器，在对 $P(\sin^2\omega/2)$ 做因式分解时，应使 Q 和 $\tilde{Q}$ 有大致相同的次数。不过这样分解的方式很有限，对 $L=(K+\tilde{K})/2<4$，惟一解为 $Q(\omega)=1$。对 $L=4$，有一个非平凡因式分解；对 $L=5$，有两个分解。表 5.1 列出了 Daubichies 等人计算的长度近似相等的滤波器系数[2]。Matlab 中的‘bior 4.4’，‘bior 5.5’及‘bior6.8’都是滤波器长度近似相等的样条双正交小波，图 5.4 画出了‘bior4.4’的尺度函数和小波，其中(a)和(b)分别为分解尺度函数和小波，而(c)和(d)分别为重构尺度函数和小波。可以明显地看到，这些对偶函数很相似，这意味着这些双正交小波很接近正交。这些双正交小波经常用于图像压缩，近似正交性保证良好的数值稳定性。

表 5.1　长度最接近的双正交滤波器

$K,\tilde{K}$	n	$h(n)$	$\tilde{h}(n)$
$K=4$, $\tilde{K}=4$	0	0.78848561640637	0.85269867800889
	-1,1	0.41809227322204	0.37740285561283
	-2,2	-0.04068941760920	-0.11062440441844
	-3,3	-0.06453888262876	-0.02384946501956
	-4,4	0	0.03782845554969
$K=5$, $\tilde{K}=5$	0	0.89950610974856	0.73666018142821
	-1,1	0.47680326579848	0.34560528195603
	-2,2	-0.09350469740094	-0.05446378846824
	-3,3	-0.13670658466433	0.00794810863724
	-4,4	-0.00269496688011	0.03968708834741
	-5,5	0.01345670945912	0

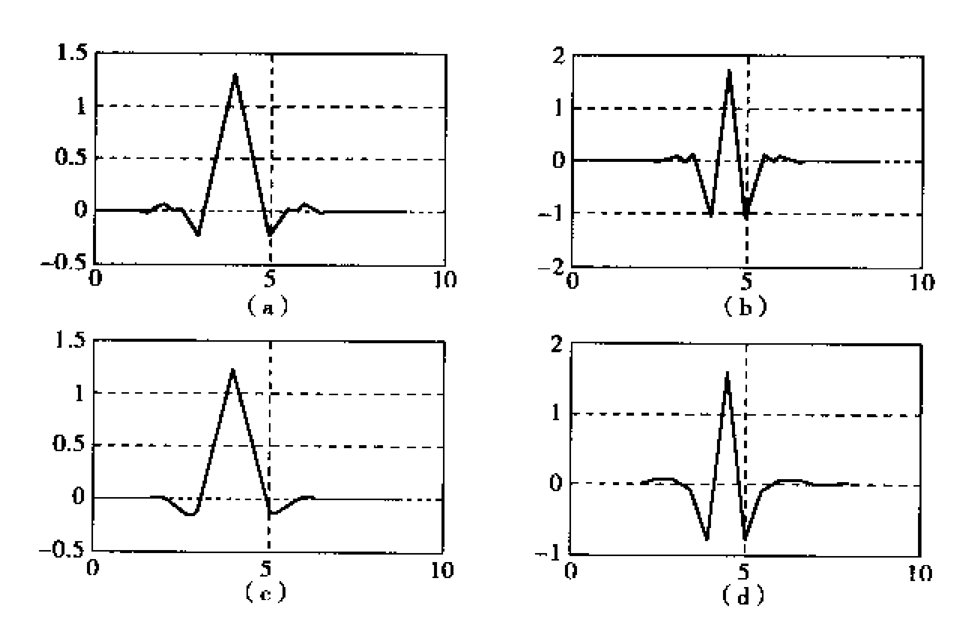


图 5.4　'bior4.4'双正交小波及尺度函数

在双正交时，Parseval 定理不再成立，也就是说，系数的范数不再等于被展开的函数的范数，这是双正交的主要缺点之一。人们已做出很多努力设计近似正交的小波基，使得范数接近相等。在随机信号处理时，经正交变换后高斯白噪声仍是白色噪声。而当变换不是正交时，噪声变得具有一定的相关性，即成为有色噪声。这样，当双正交小波用于信号检测和估计时，应当适当地调整算法以较好地清除有色噪声的影响。

5.5　二维双正交小波多分辨率分析

小波变换在图像数据压缩中的应用是一个非常活跃的研究领域。由于双正交小波的特有性质，有可能在得到高压缩比的同时，又能保持较好的图像质量。将一维情况直接推广到二

维，得到类似3.5节的快速算法，使图像的编码和解码得以高效实现。

可将一维双正交小波基直接推广得到 $L^2(R^2)$ 的可分离双正交小波基。令 ϕ,ψ 和 $\tilde{\phi},\tilde{\psi}$ 是生成 $L^2(R)$ 的双正交小波基的两对互为对偶的尺度函数和小波，如定义

$$\psi^1(x,y)=\phi(x)\psi(y),\psi^2(x,y)=\psi(x)\phi(y),\psi^3(x,y)=\psi(x)\psi(y) \quad (5.5.1)$$

和

$$\tilde{\psi}^1(x,y)=\tilde{\phi}(x)\tilde{\psi}(y),\tilde{\psi}^2(x,y)=\tilde{\psi}(x)\tilde{\phi}(y),\tilde{\psi}^3(x,y)=\tilde{\psi}(x)\tilde{\psi}(y) \quad (5.5.2)$$

则可证明，它们二进伸缩平移形成的函数族都是 $L^2(R^2)$ 的 Riesz 基。

二维双正交小波多分辨率分析的 Mallat 快速算法与3.5节基本一样，只是分解和重构要使用不同的滤波器。



原始图像



压缩图像

图5.5 双正交小波图像压缩

例5.3 双正交小波图像压缩。程序如下：

```
clear all
load woman;
level = 2;wname = 'bior4.4';
[C,S] = wavedec2(X,level,wname);
[thr,sorh,keepapp] = ddencmp('cmp','wv',X);
[Xcomp,CXC,LXC,PERF0,PERFL2] = wdencmp('gbl',C,S,wname,level,thr,sorh,keep-
app);
colormap(map);
subplot(121); image(X); title('Original Image');
subplot(122); image(Xcomp); title('Compressed Image');
```

从例3.7应该看到，图像经过小波分解之后，只是在图像的细节（边缘和纹理）处出现少量大幅值的小波系数。选取适当阈值，将小于该阈值的小波系数置0，就可以大大压缩数据，由经过数据压缩的系数重构的图像仍然有很好的质量，如图5.5所示。关于图像数据压缩，将在第8章详细讨论。在上面程序中，用函数 ddencmp 来选取阈值，并确定对系数取阈值的方式（硬阈值方式或软阈值方式）以及是否对逼近系数取阈值。一般来说，逼近保留了原始图像的概貌，应该予以保留。函数 wdencmp 用来对图像进行压缩。程序运行结果表明：所取阈值为9.5，取阈值的方式为硬阈值方式，对逼近系数不取阈值予以保留；压缩之后只保留了

47.9046%的系数,但保留了99.9293%的信号能量。

5.6　提升小波

提升可以理解为对双正交滤波器的改进。用提升方法构造的小波称为第二代小波。小波函数 $\psi_{j,n}(t)$ 通常定义为一个属于 $L^2(R)$ 的母小波的二进伸缩和平移

$$\psi_{j,n}(t) = 2^{\frac{j}{2}}\psi(2^j t - n)$$

这样的小波称为第一代小波。第二代小波并不必须是彼此的平移和伸缩,但保留了第一代小波的如下性质:

(1):构成 $L^2(R)$ 的 Riesz 基,而且还构成许多函数空间 F 如 Lebesgue、Lipschitz、Sobolev 和 Besov 空间的无条件基。函数空间中的任意一个函数 f 可以表示为 $f = \sum_{j,n}\gamma_{j,n}\psi_{j,n}$,它按 F 的范数无条件收敛。小波系数 $\gamma_{j,n} = \tilde{\psi}_{j,n}(f)$。小波是正交的,或者其对偶(双正交)小波是已知的。

(2):小波及其对偶在时域和频域都是局部化的。有些小波甚至是紧支的。

(3):小波可纳入多分辨率分析的框架。这导致了快速小波变换算法。

构造第一代小波可分为3步:第1步构造可用于快速小波变换的滤波器组,更确切地说,是寻找确保性质(3)能得到满足的多项式;第2步证明与滤波器组关联的小波是局部化的(性质2),且它们构成适当函数空间的基(性质1);第3步检查基函数的平滑性(性质2)。

Sweldens[16]分析推广了多分辨率分析的若干概念,提出了用提升方案构造小波的新途径。用提升方案构造的小波称为第二代小波。对第二代小波而言,Mallat 算法转换为对信号的偶数和奇数样本序列的逐次提升和对偶提升,称为快速多相位变换算法,它不仅减少了计算量,而且可以实现同址计算,大大减少了存储器开销。此外,第二代小波还具有许多独特的性质,从而获得了更广阔的应用领域。

5.6.1　提升的基本原理[3]

这一节介绍如何用提升方案改进 $L^2(R)$ 空间的双正交小波基。

定理5.2指出,构造紧支双正交小波基的双正交滤波器 $(h, \tilde{h}, g, \tilde{g})$ 满足

$$H^*(\omega)\tilde{H}(\omega) + H^*(\omega+\pi)\tilde{H}(\omega+\pi) = 2 \tag{5.6.1}$$

和

$$G(\omega) = e^{-i\omega}\tilde{H}^*(\omega+\pi), \tilde{G}(\omega) = e^{-i\omega}H^*(\omega+\pi) \tag{5.6.2}$$

称 h 和 $\tilde{h}$ 是对偶的。

在用提升方案构造小波时,从非常简单的滤波器 $(h_o, \tilde{h}_o, g_o, \tilde{g}_o)$ 出发,逐步地改善其性质得到提升后的滤波器 $(h, \tilde{h}_o, g_o, \tilde{g})$。也就是说,提升过程只是把 h_o 改变成 h,相应的把 $\tilde{g}_o$ 改变成 $\tilde{g}$,而 $\tilde{h}_o$ 和 g_o 并未改变。对偶提升则将 $(h_o, \tilde{h}_o, g_o, \tilde{g}_o)$ 提升为 $(h_o, \tilde{h}, g, \tilde{g}_o)$。可见提升和对偶提升过程主要是改变 h_o 或 $\tilde{h}_o$。

经过提升之后，h 与 $\tilde{h}_o$ 应该是对偶的。下面的定理阐明了与 $\tilde{h}_o$ 对偶的所有 FIR 滤波器 h 应满足的充分必要条件。

定理 5.5 设 h_o 和 $\tilde{h}_o$ 是对偶的 FIR 滤波器。FIR 滤波器 h 和 $\tilde{h}_o$ 对偶，当且仅当存在一个 FIR 滤波器 l，使得

$$H(\omega) = H_o(\omega) + e^{-i\omega}\tilde{H}_o^*(\omega+\pi)L^*(2\omega) \tag{5.6.3}$$

很容易验证，h 和 $\tilde{h}_o$ 仍然满足(5.6.1)式所示的双正交方程，即

$$H^*(\omega)\tilde{H}_o(\omega) + H^*(\omega+\pi)\tilde{H}_o(\omega+\pi) = 2$$

但上述定理意味着，由于 l 的存在，h_o 被提升为 h。提升并没有改变 $\tilde{h}_o$ 和 g_o，所以由(5.6.2)式有 $G_o(\omega) = e^{-i\omega}\tilde{H}_o^*(\omega+\pi)$，将它代入(5.6.3)式，有

$$H(\omega) = H_o(\omega) + G_o(\omega)L^*(2\omega) \tag{5.6.4}$$

相应的，$\tilde{g}_o$ 被提升为 $\tilde{g}$

$$\tilde{G}(\omega) = e^{-i\omega}H^*(\omega+\pi) = \tilde{G}_o(\omega) - \tilde{H}_o(\omega)L(2\omega) \tag{5.6.5}$$

也就是说，在老双正交滤波器($h_o, \tilde{h}_o, g_o, \tilde{g}_o$)的基础上，由于使用了 l，可以构造新的双正交滤波器($h, \tilde{h}_o, g_o, \tilde{g}$)，其性质得到了提升。如 l 是实系数的，对(5.6.4)式和(5.6.5)式做傅立叶变换的反变换可得

$$h(n) = h_o(n) + \sum_{k=-\infty}^{+\infty} g_o(n-2k)l(-k) \tag{5.6.6}$$

$$\tilde{g}(n) = \tilde{g}_o(n) - \sum_{k=-\infty}^{+\infty} \tilde{h}_o(n-2k)l(k) \tag{5.6.7}$$

提升并未改变 $\tilde{h}_o$，所以 $\tilde{\phi}_o$ 也没有改变。从提升双正交滤波器($h, \tilde{h}_o, g_o, \tilde{g}$)，通过二尺度方程，可以构造一组新的尺度函数和小波($\phi, \psi, \tilde{\phi}_o, \tilde{\psi}$)，其中

$$\phi(t) = \sqrt{2}\sum_{k=-\infty}^{+\infty} h_o(k)\phi(2t-k) + \sum_{k=-\infty}^{+\infty} l(-k)\psi(t-k) \tag{5.6.8}$$

$$\psi(t) = \sqrt{2}\sum_{k=-\infty}^{+\infty} g_o(k)\phi(2t-k) \tag{5.6.9}$$

$$\tilde{\psi}(t) = \tilde{\psi}_o(t) - \sum_{k=-\infty}^{+\infty} l(k)\tilde{\phi}_o(t-k) \tag{5.6.10}$$

上列 3 式中最值得注意的是(5.6.10)式，它的右边是在同一尺度上的老的小波和尺度函数，这意味着可以通过提升滤波器来改善 $\tilde{\psi}$ 的性质。由于 $\tilde{\psi}$ 的消失矩取决于 $H(\omega)$ 在 $\omega=\pi$ 处的零点重数，也即 $\tilde{G}(\omega)$ 在 $\omega=0$ 处的零点重数。所以可以这样来设计提升方案，在 l 最短的前提下，使 $H(\omega)$ 在 $\omega=\pi$ 处有最多的零点重数，也就是使 $\tilde{G}(\omega)$ 在 $\omega=0$ 处有最多的零点重数，从而提高 $\tilde{\psi}$ 的消失矩，与此同时，也改善了 ϕ 和 ψ 的正则性。在提升过程中，虽然由于 ϕ 的改变也改变了 ψ，但是由于 g_o 并未改变，所以 ψ 的消失矩也不会因提升而得到提高。

定理 5.6(Sweldens) 设($\phi_o, \psi_o, \tilde{\phi}_o, \tilde{\psi}_o$)是与($h_o, \tilde{h}_o, g_o, \tilde{g}_o$)相关联的一组紧支双正交

尺度函数和小波，l 是一有限序列，则与$(h,\tilde{h}_o,g_o,\tilde{g})$相关联的一组新的尺度函数和小波$(\phi,\psi,\tilde{\phi}_o,\tilde{\psi})$在形式上是双正交的。

所谓在形式上双正交，是指$(\phi,\psi,\tilde{\phi}_o,\tilde{\psi})$和$(h,\tilde{h}_o,g_o,\tilde{g})$满足二尺度方程，并不一定要求选择 l 使得提升滤波器 h 满足定理 5.4 的条件从而生成能量有限的尺度函数和小波。如果这些函数是能量有限的，那么提升后的小波$\{\psi_{j,n}\}_{j,n\in Z}$和$\{\tilde{\psi}_{j,n}\}_{j,n\in Z}$构成 $L^2(R)$的双正交小波基。

提升提高了 $\tilde{\psi}$ 的性能，为改善 ψ 的性质，可以使用对偶提升。对偶提升由滤波器 $\tilde{l}(k)$ 实现，它改变 $\tilde{h}_o$，从而也改变 g，如下列两式所示：

$$\tilde{h}(n)=\tilde{h}_o(n)+\sum_{k=-\infty}^{+\infty}\tilde{g}_o(n-2k)\tilde{l}(k) \tag{5.6.11}$$

$$g(n)=g_o(n)-\sum_{k=-\infty}^{+\infty}h_o(n-2k)\tilde{l}(-k) \tag{5.6.12}$$

对(5.6.11)式两边作傅立叶变换，有

$$\tilde{H}(\omega)=\tilde{H}_o(\omega)+e^{-i\omega}H_o^*(\omega+\pi)\tilde{L}(2\omega)$$

仍然很容易证明 $\tilde{h}$ 和 h_o 满足(5.6.1)式所示的双正交方程。对偶提升把 $\tilde{h}_o$ 改变成 $\tilde{h}$，所以也相应的把 g_o 改变成 g，且

$$G(\omega)=e^{-i\omega}\tilde{H}^*(\omega+\pi)=G_o(\omega)-H_o(\omega)\tilde{L}^*(2\omega)$$

而它正是(5.6.12)式两边作傅立叶变换的结果。

对偶提升没有改变 h_o，所以 ϕ_o 也不会改变。将对偶提升后的$(h_o,\tilde{h},g,\tilde{g}_o)$代入二尺度方程便可得到对偶提升后的尺度函数和小波$(\phi_o,\psi,\tilde{\phi},\tilde{\psi})$，其中

$$\tilde{\phi}(t)=\sqrt{2}\sum_{k=-\infty}^{+\infty}\tilde{h}_o(k)\tilde{\phi}(2t-k)+\sum_{k=-\infty}^{+\infty}\tilde{l}(k)\tilde{\psi}(t-k) \tag{5.6.13}$$

$$\tilde{\psi}(t)=\sqrt{2}\sum_{k=-\infty}^{+\infty}\tilde{g}_o(k)\tilde{\phi}(2t-k) \tag{5.6.14}$$

$$\psi(t)=\psi_o(t)-\sum_{k=-\infty}^{+\infty}\tilde{l}(-k)\phi_o(t-k) \tag{5.6.15}$$

类似的，从(5.6.15)式可以看出通过对偶提升确实可以提高 ψ 的消失矩，并改善 $\tilde{\phi}$ 和 $\tilde{\psi}$ 的正则性。不断进行提升和对偶提升可以改进 ψ 和 $\tilde{\psi}$ 的消失矩和正则性。

经过提升和对偶提升之后老滤波器$(h_o,\tilde{h}_o,g_o,\tilde{g}_o)$变为$(h,\tilde{h},g,\tilde{g})$，但在实现双正交小波多分辨率分析的 Mallat 算法时并不需要计算出提升后的滤波器，而是用老滤波器及提升和对偶提升滤波器 l 和 $\tilde{l}$ 来计算。设原始信号为 a_j，用老滤波器分解为

$$a_{j+1}^o(n)=\sum_{k=-\infty}^{+\infty}h_o(k-2n)a_j(k)=a_j*\bar{h}_o(2n) \tag{5.6.16}$$

和

$$d_{j+1}^o(n)=\sum_{k=-\infty}^{+\infty}g_o(k-2n)a_j(k)=a_j*\bar{g}_o(2n) \tag{5.6.17}$$

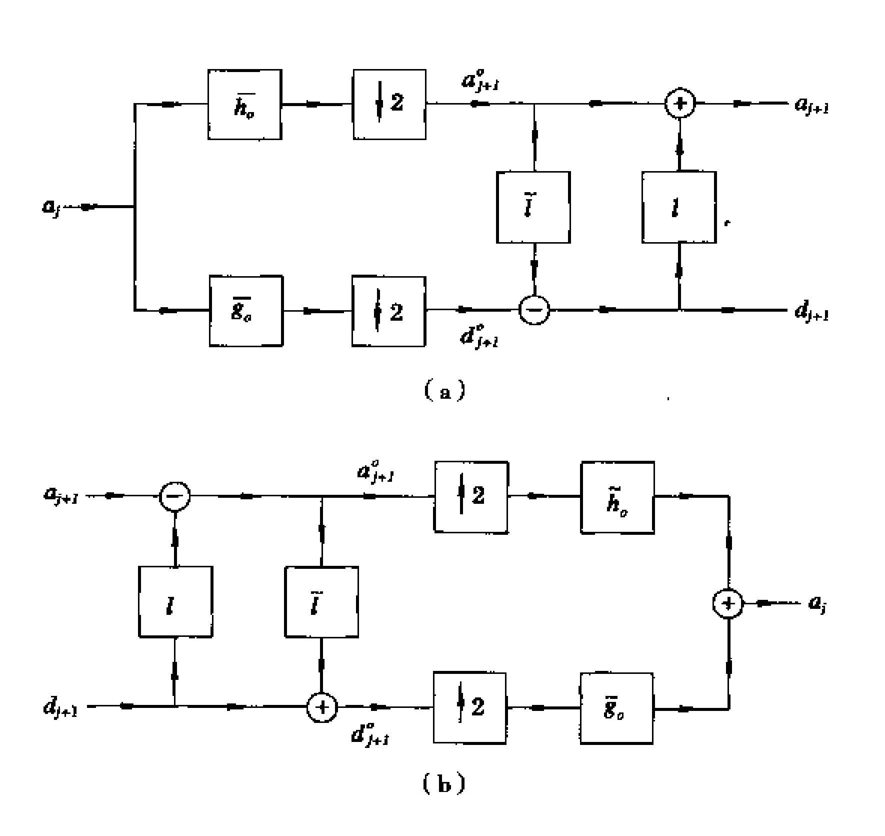


图 5.6 提升加对偶提升的多分辨率算法

(a)分解算法 (b)重构算法

先考虑对偶提升的影响，由于对偶提升并未改变 h_o，所以下 1 级的逼近分量仍然是 a^o_{j+1}，但细节分量改变为

$$d_{j+1}(n) = \sum_{k=-\infty}^{+\infty} g(k-2n)a_j(k) = a_j * \bar{g}(2n)$$

将(5.6.12)式代入上式，就得到

$$d_{j+1}(n) = d^o_{j+1}(n) - \sum_{m=-\infty}^{+\infty} a^o_{j+1}(n-m)\tilde{l}(m) = d^o_{j+1}(n) - a^o_{j+1} * \tilde{l}(n) \quad (5.6.18)$$

也就是说，经过对偶提升，细节分量修改为老细节分量加上 1 个提升分量。在对偶提升的基础上，再进行提升，则并不改变细节分量，而只是将老的逼近分量修改为

$$a_{j+1}(n) = a^o_{j+1}(n) + \sum_{m=-\infty}^{+\infty} d_{j+1}(n-m)l(m) = a^o_{j+1}(n) + d_{j+1} * l(n) \quad (5.6.19)$$

上列两式就是对偶提升加提升的多分辨率分解算法，如图 5.6(a)所示。将上列两式的提升和对偶提升分量移到等式的左边，就可以得到重构算法，如图 5.6(b)所示。

5.6.2 惰性小波、提升方案快速多相位小波变换算法

提升方案的优越性还在于，甚至可以从非常简单平凡的小波出发而设计出具有一定性质的小波。惰性小波就是非常简单平凡的小波。

惰性滤波器 $\tilde{h}_o(n) = h_o(n) = \delta(n)$，$\tilde{g}_o(n) = g_o(n) = \delta(n-1)$，它们的傅立叶变换是

$$\tilde{H}_o(\omega) = H_o(\omega) = 1, \qquad \tilde{G}_o(\omega) = G_o(\omega) = e^{-i\omega}$$

惰性滤波器显然满足双正交条件(5.6.1)式和(5.6.2)式。惰性滤波器没有滤波，只是将信号

的偶数和奇数样本分离，这也称为多相位分解。与惰性滤波器关联的惰性尺度函数和小波是冲击函数 $\tilde{\phi}_o(t)=\phi_o(t)=\delta(t)$ 和 $\tilde{\psi}_o(t)=\psi_o(t)=\delta(t-1/2)$，它们不属于 $L^2(R)$，理由是 $\tilde{G}_o(\omega)$ 和 $G_o(\omega)$ 在 $\omega=0$ 处不等于零。但是通过适当的提升，这些小波可以变成能量有限的函数。

定理 5.7(Daubechies,Sweldens)　任何双正交滤波器$(h,g,\tilde{h},\tilde{g})$可以通过对惰性滤波器不断做提升和对偶提升而生成(至多相差一个移位和常数因子)。

上述定理引出的一个非常重要的结论是可以用提升和对偶提升实现 Mallat 算法，将提升滤波器表示为 u，从而可将(5.6.6)式和(5.6.7)式写成

$$h(n)=h_o(n)+\sum_{k=-\infty}^{+\infty}g_o(n-2k)u(-k) \tag{5.6.20}$$

$$\tilde{g}(n)=\tilde{g}_o(n)-\sum_{k=-\infty}^{+\infty}\tilde{h}_o(n-2k)u(k) \tag{5.6.21}$$

将提升滤波器表示为 p，从而可将对偶提升(5.6.11)式和(5.6.12)式写成

$$\tilde{h}(n)=\tilde{h}_o(n)+\sum_{k=-\infty}^{+\infty}\tilde{g}_o(n-2k)p(k) \tag{5.6.22}$$

$$g(n)=g_o(n)-\sum_{k=-\infty}^{+\infty}h_o(n-2k)p(-k) \tag{5.6.23}$$

对于惰性滤波器

$$a_{j+1}^o(n)=a_j(2n),d_{j+1}^o(n)=a_j(2n+1) \tag{5.6.24}$$

它只是将输入信号分裂为偶数和奇数样本序列，称为信号的多相位分解，对多相位分量先对偶提升再进行提升，则有

$$d_{j+1}(n)=d_{j+1}^o(n)-a_{j+1}^o*p(n) \tag{5.6.25}$$

$$a_{j+1}(n)=a_{j+1}^o(n)+d_{j+1}*u(n) \tag{5.6.26}$$

上列两式就是对偶提升加提升的多相位分解算法，如图5.7(a)所示。同样将上列两式的提升和对偶提升分量移到等式的左边，就可以得到重构算法，如图5.7(b)所示。

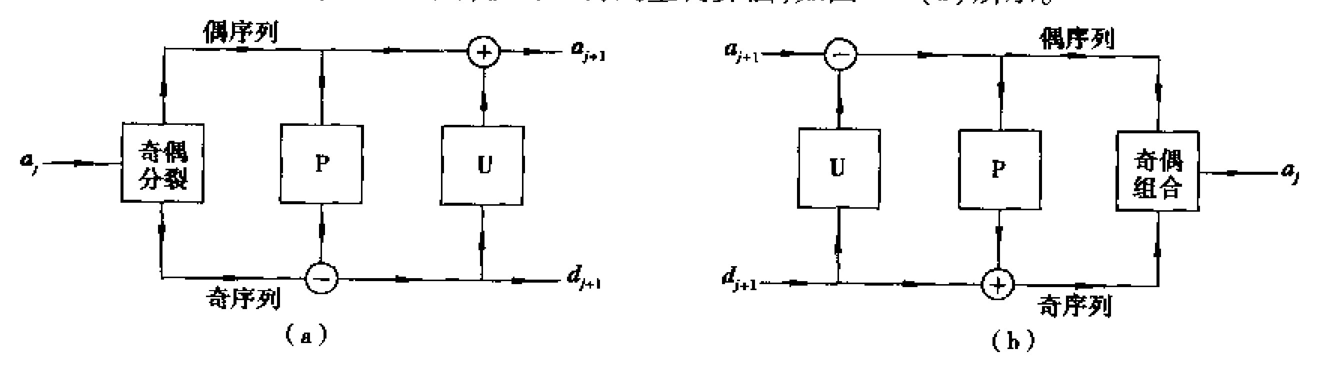


图5.7　提升方案快速多相位变换算法

(a)分解算法　(b)重构算法

在分解过程中，首先将 a_j 分解为 a_{j+1}^o 和 d_{j+1}^o；对偶提升由 a_{j+1}^o 和 d_{j+1}^o 计算 d_{j+1}，由于提升不需要 d_{j+1}^o，所以在实际计算过程中可以用 d_{j+1} 来刷新 d_{j+1}^o 的存储空间；提升由 a_{j+1}^o 和 d_{j+1} 计算 a_{j+1}，也可以用 a_{j+1} 来刷新 a_{j+1}^o 的存储空间。这就是所谓的同址计算，它可以大大地减少存储器开销。此外，用提升方案实现小波变换的计算量是直接用 Mallat 算法的计算量除以某个因

子(最大为2)。

由上述分析可以看到,提升方案把小波变换分为分裂、预测和更新3个阶段:

(1)**分裂**(split)

分裂过程就是惰性小波变换过程。此过程仅仅是将数据集合 a_j 分裂成它的偶数样本子集和奇数样本子集

$$\mathrm{split}(a_j) = (a_{j+1}, d_{j+1}) \tag{5.6.27}$$

(2)**预测**(predict)

预测过程就是对偶提升过程。由于两个子集之间总是存在一定的相关性,所以对偶提升分量可以看成偶序列对奇序列的预测值,于是对偶提升过程实际上就是用预测误差来代替奇序列

$$d_{j+1} \leftarrow d_{j+1} - P(a_{j+1}) \tag{5.6.28}$$

预测误差又称为细节系数或小波系数,小波系数越小,预测越精确。

(3)**更新**(update)

更新过程就是提升过程。通过数据集合 d_{j+1} 来更新数据集合 a_{j+1}

$$a_{j+1} \leftarrow a_{j+1} + U(d_{j+1}) \tag{5.6.29}$$

更新的出发点是找到一个比 a_{j+1} 更好的数据子集,使之保持原始数据的一些特性,如能量、均值、消失矩等。更新过程用提升算子作用于细节系数并叠加到偶数样本得到近似信号。

重构算法由如下3个步骤组成

$$a_{j+1} \leftarrow a_{j+1} - U(d_{j+1}) \tag{5.6.30}$$

$$d_{j+1} \leftarrow d_{j+1} + P(a_{j+1}) \tag{5.6.31}$$

$$a_j = \mathrm{merge}(a_{j+1}, d_{j+1}) \tag{5.6.32}$$

同样,重构算法也可以用同址计算实现。

5.6.3 任意区域上的提升小波[16]

由于提升小波并不必须是彼此的平移和伸缩,所以用提升方法可以构造非平移不变信号空间的小波基。放弃平移和伸缩意味着不再用傅立叶变换作为构造小波的工具。

设 $L_2(D)$ 是定义在 R^n 的任意区域 $D \subset R^n$ 上的任意函数空间,为了定义它的多分辨率逼近,引入更广义的指标集 $\{G_j\}_{j\in Z}$,且 $G_j \subset G_{j+1}$,G_j 的指标为 k。C_j 为 G_j 在 G_{j+1} 中的补,它是所有属于 G_{j+1} 而又不属于 G_j 的元素的集合,C_j 的指标表示为 m。G_j 也可以理解为区域 D 上的采样网格[3],例如,$D=[0,1]$ 时的二进均匀采样网格 G_j 是 $\{2^{-j}n\}_{0\leqslant n\leqslant 2^j}$,当 $j=2$ 时的网格节点如图5.8中的小圆圈所示,这些网格节点分别对应0/4、1/4、2/4、3/4、4/4,或写成0/8、2/8、4/8、6/8、8/8,它们也就是指标集中的 k 值。这时 G_j 在 G_{j+1} 中的补 C_j 聚集了所有属于 G_{j+1} 而又不属于 G_j 的网格节点,即 $\{2^{-(j+1)}(2n+1)\}_{0\leqslant n<2^j}$,当 $j=2$ 时 C_j 如图5.8中的短划线所示,它们分别对应1/8、3/8、5/8、7/8,也就是 m 的值。

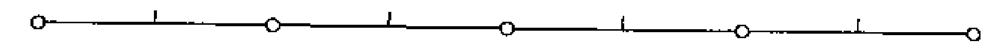


图5.8 [0,1]区间的二进均匀网格及其补网格

Sweldens 在用提升方案构造第二代小波时,从 $L_2(D)$ 的多分辨率分析入手。如嵌套闭子空间序列 $\{V_j\}_{j\in Z}$ 满足与定义3.1类似的某些性质,则它们构成 $L_2(D)$ 的多分辨率逼近,尺度

函数$\{\phi_{j,k}|k\in G_j\}$构成V_j的Riesz基。但并不要求$\phi_{j,k}$是某个尺度函数的平移和伸缩。此外，还将第一代小波在实轴上的二进区间推广为区域D的划分，划分的每一个子集与一个尺度函数$\phi_{j,k}$关联。$\{\tilde{V}_j\}_{j\in Z}$表示$L_2(D)$的对偶多分辨率逼近，其中对偶尺度函数$\tilde{\phi}_{j,k}$构成$\tilde{V}_j$的Riesz基。$\phi_{j,k}$和$\tilde{\phi}_{j,k}$在如下意义上是双正交的

$$\langle\phi_{j,k},\tilde{\phi}_{j,k'}\rangle=\delta_{k,k'}\quad 对\quad k,k'\in G_j$$

对任意$f\in L_2$，定义系数$\lambda_{j,k}=\langle f,\tilde{\phi}_{j,k}\rangle$，则$f$在$V_j$上的投影

$$P_jf=\sum_{k\in G_j}\lambda_{j,k}\phi_{j,k}$$

是在尺度j时对f的逼近，且满足

$$\lim_{j\to\infty}\|f-P_jf\|=0$$

与第一代小波类似，仍然通过设计滤波器来构造尺度函数和小波。由于$V_j\subset V_{j+1}$，所以$\phi_{j,k}$可以表示为更精细尺度上的尺度函数$\phi_{j+1,l}$的线性组合

$$\phi_{j,k}=\sum_{l\in G_{j+1}}h_{j,k,l}\phi_{j+1,l}\tag{5.6.33}$$

其中，实数集合$\{h_{j,k,l}|k\in G_j,l\in G_{j+1}\}$称为滤波器。务必注意，第二代小波中的滤波器$h_{j,k,l}$一般因分辨率精度$j$不同而不同，所以并不一定具有如下平移不变性：$h_{j,k,l}=h_{l-2k}$，因此(5.6.33)式一般而言并不是第一代小波的二尺度方程，而被称为细化关系式。与第一代小波类似，通常都难以给出尺度函数的解析表达式，但仍然可以通过级联算法的迭代计算过程由滤波器定义相应的尺度函数。

第一代小波定义为V_j在V_{j+1}中的补W_j的基函数，这一概念也被推广到第二代小波，即小波被定义为　$W_j=\overline{\text{span}\{\psi_{j,m}|m\in C_j\}}$。双正交小波的含义也由如下内积关系式给出

$$\langle\psi_{j,m},\tilde{\psi}_{j',m'}\rangle=\delta_{j,j'}\delta_{m,m'}$$

对任意$f\in L_2$，定义系数$\gamma_{j,m}=\langle f,\tilde{\psi}_{j,m}\rangle$，则

$$f=\sum_{j,m}\gamma_{j,m}\psi_{j,m}$$

小波的定义意味着小波满足与(5.6.33)式相似的细化关系式

$$\psi_{j,m}=\sum_{l}g_{j,m,l}\phi_{j+1,l}\tag{5.6.34}$$

其中，$\{g_{j,m,l}|m\in C_j,l\in G_{j+1}\}$也是一个滤波器。至此，已经分别用滤波器$h_{j,k,l}$和$g_{j,m,l}$通过细化关系式(5.6.33)式和(5.6.34)式定义了尺度函数和小波。对偶尺度函数和小波与对偶滤波器$\tilde{h}_{j,k,l}$和$\tilde{g}_{j,m,l}$之间也有类似的细化关系式，而且由$\phi_{j+1,l}\in V_j\oplus W_j$，可推知

$$\phi_{j+1,l}=\sum_{k}\tilde{h}_{j,k,l}\phi_{j,k}+\sum_{m}\tilde{g}_{j,m,l}\psi_{j,m}$$

由尺度函数和小波的双正交表达式还可推知

$$\sum_{l}h_{j,k,l}\tilde{h}_{j,k',l}=\delta_{k,k'},\ \sum_{l}h_{j,k,l}\tilde{g}_{j,m,l}=0$$

$$\sum_{l}g_{j,m,l}\tilde{g}_{j,m',l}=\delta_{m,m'},\ \sum_{l}g_{j,m,l}\tilde{h}_{j,k,l}=0\tag{5.6.35}$$

满足上式的滤波器$\{h,\tilde{h},g,\tilde{g}\}$称为双正交滤波器。

小波变换的基本思想也与第一代小波相同。由初始尺度 j_0 上的原始信号序列 $\{\lambda_{j_0,k} | k \in G_{j_0}\}$ 可以逐级分解成如下信号序列集：

$$\{\lambda_{J,k} \mid k \in G_J; \gamma_{j,m} \mid J \leqslant j < j_0, m \in C_j\} \tag{5.6.36}$$

从细化关系可以得到小波变换快速分解表达式如下

$$\lambda_{j,k} = \sum_l \tilde{h}_{j,k,l} \lambda_{j+1,l} \tag{5.6.37}$$

$$\gamma_{j,m} = \sum_l \tilde{g}_{j,m,l} \lambda_{j+1,l} \tag{5.6.38}$$

类似的，小波变换快速重构表达式为

$$\lambda_{j+1,l} = \sum_k h_{j,k,l} \lambda_{j,k} + \sum_m g_{j,m,l} \gamma_{j,m} \tag{5.6.39}$$

与第一代小波不同的是，因为滤波器系数不一定具有平移不变性，所以第二代小波变换分解和重构表达式中的线性算子不一定是卷积运算。

很容易证明，双正交滤波器经过提升之后可以得到新的双正交滤波器。

定理 5.8 老双正交滤波器 $\{h^{\text{old}}, \tilde{h}^{\text{old}}, g^{\text{old}}, \tilde{g}^{\text{old}}\}$ 经过如下提升可以构成新双正交滤波器 $\{h, \tilde{h}, g, \tilde{g}\}$

$$h_{j,k,l} = h_{j,k,l}^{\text{old}}$$

$$\tilde{h}_{j,k,l} = \tilde{h}_{j,k,l}^{\text{old}} + \sum_m s_{j,k,m} \tilde{g}_{j,m,l}^{\text{old}}$$

$$g_{j,m,l} = g_{j,m,l}^{\text{old}} - \sum_k s_{j,k,m} h_{j,k,l}^{\text{old}}$$

$$\tilde{g}_{j,m,l} = \tilde{g}_{j,m,l}^{\text{old}}$$

经过提升之后，因为 h 没改变，所以尺度函数 ϕ 也不会改变。而 $\tilde{h}$ 和 g 的改变使得对偶尺度函数 $\tilde{\phi}$ 和小波 ψ 也要改变。虽然 $\tilde{g}$ 没改变，但由于 $\tilde{\phi}$ 改变了，所以对偶小波也要改变。从细化关系可得下列各式

$$\phi_{j,k} = \phi_{j,k}^{\text{old}} \tag{5.6.40}$$

$$\tilde{\phi}_{j,k} = \sum_l \tilde{h}_{j,k,l}^{\text{old}} \tilde{\phi}_{j+1,l} + \sum_m s_{j,k,m} \tilde{\psi}_{j,m} \tag{5.6.41}$$

$$\psi_{j,m} = \psi_{j,m}^{\text{old}} - \sum_k s_{j,k,m} \phi_{j,k}^{\text{old}} \tag{5.6.42}$$

$$\tilde{\psi}_{j,m} = \sum_l \tilde{g}_{j,m,l}^{\text{old}} \tilde{\phi}_{j+1,l} \tag{5.6.43}$$

应特别注意(5.6.42)式，它的右边都是老的尺度函数和小波，这意味着通过适当选择 $s_{j,k,m}$ 从简单的尺度函数和小波可以得到具有特定性质的小波，这正是提升的含义。该式还意味着提升方法的关键是寻找适当的提升算子 S，因为对小波性质的要求都转换成对 $s_{j,k,m}$ 的约束条件。而且通过提升方法构造小波时是在同一尺度下进行的，而不是像(5.6.34)式表示的那样，要从更精细一级尺度上的尺度函数来构造小波。最后，定理 5.8 还确保：只要老滤波器组是双正交的，则提升滤波器组也必然是双正交的。

惰性滤波器 $h_{j,k,l} = \delta_{k,l}, g_{j,m,l} = \delta_{m,l}$，它们的对偶滤波器也是惰性的，相应的尺度函数和小波是在指标集上的冲击函数。通过不断的提升和对偶提升也可以从惰性滤波器得到具有一定

性质的双正交滤波器，从而构造出具有特定性质的双正交小波基。惰性滤波器只不过是把属于 G_{j+1} 的样本分裂成为分别属于 G_j 和 C_j 的两个样本集。对在嵌套采样网格上构造的惰性小波作提升，可算出流形或 R^n 的有界区域上的双正交小波基。曲面上的提升小波在计算机图形学中有许多应用。在有限二维区域中，提升小波基应用于偏微分方程的数值计算。

第 6 章

信号的小波包分析

在信号的多分辨率分析中,首先将原始信号 a_0分解为低频部分 a_1和高频部分 d_1,再将 a_1进一步分解为 a_2和 d_2,这样逐级分解下去,如图 6.1 所示。可以看到,在信号的多分辨率分析中,总是对信号的低频部分进行分解,而没对高频部分进行进一步的分解,所以信号的多分辨率分析是一个不完整的树状结构。

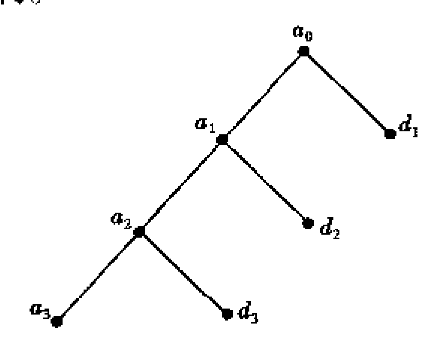


图 6.1　信号的多分辨率分析

对信号进行小波包分析时,不仅对信号的低频部分进行分解,同时也对高频部分进行分解。所以,信号的小波包分析是一个完整的树状结构,如图 6.2 所示。不难想到,小波包分析比多分辨率分析更精细。当然,其计算复杂性更高,计算量更大。我们知道,小波分析在频率域相当于等 Q 带通滤波器组,尺度越精细,带通滤波器的中心频率越高,通带越宽。也就是说,对信号的高频成分,小波分析的分辨率较差,这是小波分析的严重不足。而小波包分析对信号的高频成分也有较高分辨率,这正是小波包分析的优势。

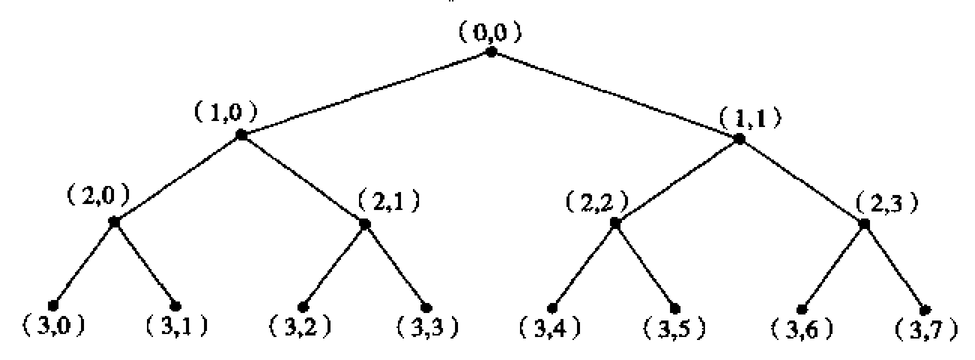


图 6.2　信号的小波包分析

在图6.2的二叉树中,圆括弧内的第一个数表示分解的尺度级次(或者说分辨率级次);第二个数是小波包参数,当它为0时,对应多分辨率分析的低频分量,当它为1时,对应多分辨率分析的高频分量。而当它为大于1的正整数时,就是新增的小波包分析。可以看出,小波包分析实际上是将多分辨率分析的高频分量进一步分解成低频部分和高频部分。

6.1 小波包

6.1.1 小波包的定义

信号的多分辨率分析实质上是两个相邻分辨率之间信号逐级分解的过程。在函数空间,分别用尺度函数 ϕ 和小波 ψ 分解出信号的低频部分和高频部分。而在序列空间,则用一对滤波器 h 和 g 对离散逼近系数进行分解。后者可用 Mallat 算法高效实现。ϕ,ψ 和与它们关联的 h,g 之间存在如下二尺度方程:

$$\phi(t) = \sqrt{2}\sum_k h(k)\phi(2t-k) \tag{6.1.1}$$

$$\psi(t) = \sqrt{2}\sum_k g(k)\phi(2t-k) \tag{6.1.2}$$

$\{2^{-j/2}\phi(2^{-j}t-k)\}_{k\in Z}$ 和 $\{2^{-j/2}\psi(2^{-j}t-k)\}_{k\in Z}$ 分别构成尺度函数空间 V_j 和小波函数空间 W_j 的正交规范基。上列两式右边的 $\phi(2t)$ 对应着尺度较小(分辨率较高)的函数空间 V_j,而左边的 $\phi(t)$ 和 $\psi(t)$ 则对应着尺度较大(分辨率较低)的函数空间 V_{j+1} 和 W_{j+1}。所以(6.1.1)式和(6.1.2)式本质上反映了函数空间的如下关系

$$V_j = V_{j+1} \oplus W_{j+1}, j \in Z \tag{6.1.3}$$

这就是多分辨率分析的函数空间分解关系式。

将二尺度方程(6.1.1)式和(6.1.2)式推广为

$$u_{2n}(t) = \sqrt{2}\sum_k h(k)u_n(2t-k) \tag{6.1.4}$$

$$u_{2n+1}(t) = \sqrt{2}\sum_k g(k)u_n(2t-k) \tag{6.1.5}$$

其中,下标 n 表示函数的序号,$n=0,1,2,\cdots$,记为 $n\in Z^+$。令 $n=0$,$u_0=\phi$,$u_1=\psi$,则上列两式就退化为(6.1.1)式和(6.1.2)式。(6.1.4)式和(6.1.5)式是一对递推关系式,首先由 $u_0=\phi$ 生成 $u_1=\psi$,就可由 $u_1=\psi$ 生成 u_2 和 u_3,再由 u_2 生成 u_4 和 u_5,由 u_3 生成 u_6 和 u_7,如此逐级递推下去,生成一个函数族 $\{u_n\}_{n\in Z^+}$。

定义6.1 小波包的定义:称由(6.1.4)式和(6.1.5)式定义的函数族 $\{u_n\}_{n\in Z^+}$ 为由 ϕ 生成的小波包,或由 h 生成的小波包。其中下标 n 为小波包参数。

小波包的生成如图6.3所示。小波包参数 $n=0$ 对应尺度函数,$n=1$ 对应小波。比较图6.3和图6.2不难发现,小波包参数 n 正好对应着图6.2中圆括弧内第二个数。由(6.1.4)式和(6.1.5)式可以明显地看到,小波包参数为偶数对应低通滤波器 h,而小波包参数为奇数对应高通滤波器 g,所以小波包分析可将多分辨率分析的高频分量进一步逐级分解为两部分。

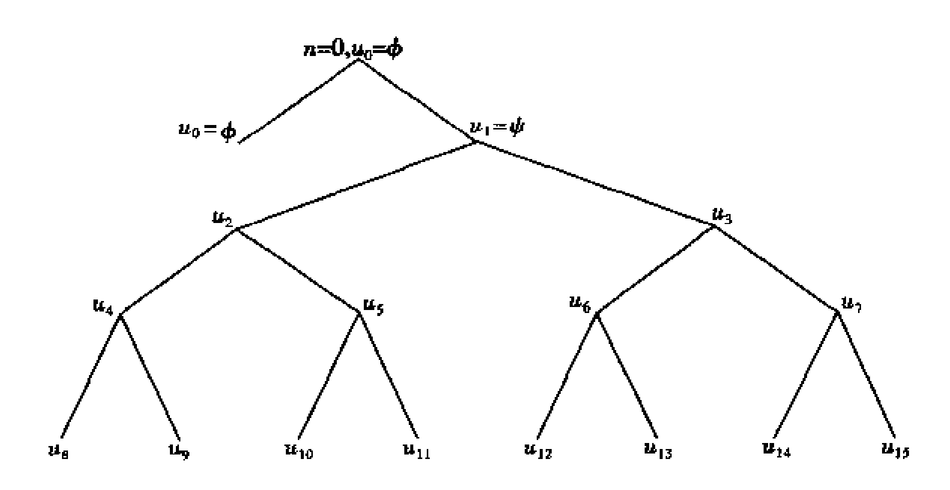


图 6.3　小波包的生成

6.1.2　小波包的性质

对(6.1.1)式、(6.1.2)式及(6.1.4)式、(6.1.5)式两边取傅立叶变换,有

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}}H\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \tag{6.1.6}$$

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}G\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \tag{6.1.7}$$

和

$$\hat{u}_{2n}(\omega) = \frac{1}{\sqrt{2}}H\left(\frac{\omega}{2}\right)\hat{u}_n\left(\frac{\omega}{2}\right) \tag{6.1.8}$$

$$\hat{u}_{2n+1}(\omega) = \frac{1}{\sqrt{2}}G\left(\frac{\omega}{2}\right)\hat{u}_n\left(\frac{\omega}{2}\right) \tag{6.1.9}$$

若令

$$H(\omega) = P_0(\omega) \tag{6.1.10}$$

$$G(\omega) = P_1(\omega) \tag{6.1.11}$$

$\lfloor n \rfloor$ = 不超过 n 的最大整数

则(6.1.6)式~(6.1.9)式可统一表示为

$$\hat{u}_n(\omega) = \frac{1}{\sqrt{2}}P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\hat{u}_{\lfloor n/2\rfloor}\left(\frac{\omega}{2}\right), n \in Z^+ \tag{6.1.12}$$

其中,ε_1 为 n 二进制表式的第 1 位。当 n 为偶数时,$\varepsilon_1=0$;当 n 为奇数时,$\varepsilon_1=1$。

由于 n 是一个非负整数,所以其二进制表式为

$$n = \sum_{j=1}^{+\infty}\varepsilon_j 2^{j-1}, \varepsilon_j = 0 \text{ 或 } 1 \tag{6.1.13}$$

反复用(6.1.12)式,可得到 $\hat{u}_n(\omega)$ 的一个无穷乘积极限形式。

定理 6.1　由 h 生成的小波包 $\{u_n\}_{n\in Z^+}$ 的傅立叶变换为

$$\hat{u}_n(\omega) = \prod_{j=1}^{+\infty}\left[\frac{1}{\sqrt{2}}P_{\varepsilon_j}(2^{-j}\omega)\right] \tag{6.1.14}$$

证明　首先证明 $n=0$ 和 1 时,上式成立。这时,$\lfloor n/2\rfloor=0$,由(6.1.12)式,有

$$\hat{u}_n(\omega) = \frac{1}{\sqrt{2}} P_{\varepsilon_1}(2^{-1}\omega)\hat{u}_0(2^{-1}\omega)$$

再反复用(6.1.6)式，可得

$$\hat{u}_n(\omega) = \frac{1}{\sqrt{2}} P_{\varepsilon_1}(2^{-1}\omega) \prod_{j=2}^{+\infty} \left[\frac{1}{\sqrt{2}} P_0(2^{-j}\omega)\right] = \prod_{j=1}^{+\infty} \left[\frac{1}{\sqrt{2}} P_{\varepsilon_j}(2^{-j}\omega)\right]$$

然后，用归纳法证明(6.1.14)式对任意的 n 成立。假定当 $0 \leqslant n < 2^J$ 时(6.1.14)式成立，考虑当 $2^J \leqslant n < 2^{J+1}$ 时的情形。这时，有

$$n = \sum_{j=1}^{J+1} \varepsilon_j 2^{j-1}, \varepsilon_{J+1} = 1$$

而 $\lfloor n/2 \rfloor$ 实际上是 n 除以 2 取商的整数部分，也就是将 n 的二进制表示中小数点左移 1 位后取整，即

$$\lfloor n/2 \rfloor = \sum_{j=1}^{J} \varepsilon_{j+1} 2^{j-1}$$

由于 $0 < \lfloor n/2 \rfloor < 2^J$，故由(6.1.12)式和归纳假设有

$$\begin{aligned}
\hat{u}_n(\omega) &= \frac{1}{\sqrt{2}} P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\hat{u}_{\lfloor n/2 \rfloor}\left(\frac{\omega}{2}\right) \\
&= \frac{1}{\sqrt{2}} P_{\varepsilon_1}(2^{-1}\omega) \prod_{j=1}^{+\infty} \left[\frac{1}{\sqrt{2}} P_{\varepsilon_{j+1}}(2^{-(j+1)}\omega)\right] \\
&= \frac{1}{\sqrt{2}} P_{\varepsilon_1}(2^{-1}\omega) \prod_{j=2}^{+\infty} \left[\frac{1}{\sqrt{2}} P_{\varepsilon_j}(2^{-j}\omega)\right] \\
&= \prod_{j=1}^{+\infty} \left[\frac{1}{\sqrt{2}} P_{\varepsilon_j}(2^{-j}\omega)\right]
\end{aligned}$$

证毕。由正交镜像滤波器 h,g 的正交性可以导出小波包具有如下正交性质。

定理 6.2　设 $\{u_n\}_{n \in Z^+}$ 是由 h 生成的小波包，则

$$\langle u_n(t-k), u_n(t-l) \rangle = \delta_{k,l}, k,l \in Z \tag{6.1.15}$$

证明　当 $n=0$ 和 1 时，(6.1.15)式显然成立。假设当 $0 \leqslant n < 2^J$ 时(6.1.15)式成立，考虑当 $2^J \leqslant n < 2^{J+1}$ 时的情形。

为简化记号，考虑如下等价内积表达式

$$\begin{aligned}
\langle u_n(t), u_n(t-k) \rangle &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} |\hat{u}_n(\omega)|^2 e^{i\omega k} d\omega \\
&= \frac{1}{4\pi}\int_{-\infty}^{+\infty} \left|P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\right|^2 \left|\hat{u}_{\lfloor n/2 \rfloor}\left(\frac{\omega}{2}\right)\right|^2 e^{i\omega k} d\omega \\
&= \frac{1}{4\pi}\sum_{l=-\infty}^{+\infty}\int_{4\pi l}^{4\pi(l+1)} \left|P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\right|^2 \left|\hat{u}_{\lfloor n/2 \rfloor}\left(\frac{\omega}{2}\right)\right|^2 e^{i\omega k} d\omega \\
&= \frac{1}{4\pi}\sum_{l=-\infty}^{+\infty}\int_{0}^{4\pi} \left|P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\right|^2 \left|\hat{u}_{\lfloor n/2 \rfloor}\left(\frac{\omega}{2}+2\pi l\right)\right|^2 e^{i\omega k} d\omega \\
&= \frac{1}{4\pi}\int_{0}^{4\pi} e^{i\omega k} \left|P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\right|^2 \sum_{l=-\infty}^{+\infty} \left|\hat{u}_{\lfloor n/2 \rfloor}\left(\frac{\omega}{2}+2\pi l\right)\right|^2 d\omega
\end{aligned}$$

由于 $0 < \lfloor n/2 \rfloor < 2^J$，故归纳假设意味着下式成立

$$\langle u_{\lfloor n/2\rfloor}(t-k),u_{\lfloor n/2\rfloor}(t-l)\rangle=\delta_{k,l},k,l\in Z$$

它等价于

$$\sum_{l=-\infty}^{+\infty}\left|\hat{u}_{\lfloor n/2\rfloor}\left(\frac{\omega}{2}+2\pi l\right)\right|^2=1$$

所以,有

$$\begin{aligned}\langle u_n(t),u_n(t-k)\rangle&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}|\hat{u}_n(\omega)|^2\mathrm{e}^{i\omega k}\mathrm{d}\omega\\&=\frac{1}{4\pi}\int_0^{2\pi}\mathrm{e}^{i\omega k}\left|P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\right|^2\mathrm{d}\omega+\frac{1}{4\pi}\int_{2\pi}^{4\pi}\mathrm{e}^{i\omega k}\left|P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\right|^2\mathrm{d}\omega\\&=\frac{1}{4\pi}\int_0^{2\pi}\mathrm{e}^{i\omega k}\left[\left|P_{\varepsilon_1}\left(\frac{\omega}{2}\right)\right|^2+\left|P_{\varepsilon_1}\left(\frac{\omega}{2}+\pi\right)\right|^2\right]\mathrm{d}\omega\\&=\frac{1}{2\pi}\int_0^{2\pi}\mathrm{e}^{i\omega k}\mathrm{d}\omega=\delta_k\end{aligned}$$

上面用到 h,g 的正交性：$|H(\omega)|^2+|H(\omega+\pi)|^2=2$ 或 $|G(\omega)|^2+|G(\omega+\pi)|^2=2$。证毕。

上述定理表明小波包中每个函数都具有平移正交性。下面将证明 $u_{2n}(t)$ 和 $u_{2n+1}(t)$ 之间的平移正交性。证毕。

定理 6.3 设 $\{u_n\}_{n\in Z^+}$ 是由 h 生成的小波包,则

$$\langle u_{2n}(t-k),u_{2n+1}(t-l)\rangle=0,k,l\in Z \tag{6.1.16}$$

证明 当 $n=0$ 时,(6.1.16)式显然成立。这时实际上是尺度函数和小波正交,它在频域的等价表达式为 $H(\omega)G^*(\omega)+H(\omega+\pi)G^*(\omega+\pi)=0$。下面的过程和定理 6.2 的证明很相似。由(6.1.12)式有

$$\begin{aligned}&\langle u_{2n}(t-k),u_{2n+1}(t-l)\rangle\\&=\frac{1}{4\pi}\int_{-\infty}^{+\infty}P_0\left(\frac{\omega}{2}\right)P_1^*\left(\frac{\omega}{2}\right)\left|\hat{u}_n\left(\frac{\omega}{2}\right)\right|^2\mathrm{e}^{i(l-k)\omega}\mathrm{d}\omega\\&=\frac{1}{4\pi}\int_0^{4\pi}\mathrm{e}^{i(l-k)\omega}P_0\left(\frac{\omega}{2}\right)P_1^*\left(\frac{\omega}{2}\right)\sum_{q=-\infty}^{+\infty}\left|\hat{u}_n\left(\frac{\omega}{2}+2\pi q\right)\right|^2\mathrm{d}\omega\\&=\frac{1}{4\pi}\int_0^{2\pi}\mathrm{e}^{i(l-k)\omega}\left[P_0\left(\frac{\omega}{2}\right)P_1^*\left(\frac{\omega}{2}\right)+P_0\left(\frac{\omega}{2}+\pi\right)P_1^*\left(\frac{\omega}{2}+\pi\right)\right]\mathrm{d}\omega\\&=0\end{aligned}$$

上面用到 h,g 之间的正交性：$H(\omega)G^*(\omega)+H(\omega+\pi)G^*(\omega+\pi)=0$ 以及 u_n 的平移正交性。证毕。

6.1.3 紧支集小波包

可以从任意有限冲击响应正交镜像滤波器 h 和 g 生成紧支集小波包。设 h 和 g 支集为 $[0,N-1]$,那么,ϕ 和 ψ 的支集长度为 $(N-1)$。

这时,二尺度方程(6.1.4)式和(6.1.5)式可写为：

$$u_{2n}(t)=\sqrt{2}\sum_{k=0}^{N-1}h(k)u_n(2t-k) \tag{6.1.17}$$

$$u_{2n+1}(t)=\sqrt{2}\sum_{k=0}^{N-1}g(k)u_n(2t-k) \tag{6.1.18}$$

可见 u_{2n} 和 u_{2n+1} 是 u_n 按 1/2 尺度压缩、平移后加权叠加的结果；因此，所有 u_n 的支集长度均为 $(N-1)$。既然所有 u_n 都有相同的支集长度，而 n 越大又意味着压缩的次数越多，所以粗略地说，n 描述了小波包中各个波形函数的振荡频率，故又称其为频率参数。

例 6.1　Harr 小波包。

Harr 小波包由滤波器

$$h(n) = \begin{cases} 1/\sqrt{2}, & n = 0,1 \\ 0, & \text{其他} \end{cases}$$

生成。图 6.4 画出了 Harr 小波包的波形，可以清楚地看到 n 确实可以解释为振荡频率。u_0 是尺度函数 ϕ，它正是[0,1]区间的特征函数。u_1 是哈尔小波，其正负脉冲宽度均为 1/2。u_2, u_3 由 u_1 压缩平移线性叠加生成，它们的正负脉冲宽度均为 1/4，即它们具有相同的频率。u_4, u_5 由 u_2 压缩平移线性叠加生成，u_6, u_7 则由 u_3 压缩平移线性叠加生成，所以这 4 个函数的正负脉冲宽度均为 1/8，具有相同的频率。也就是说，在图 6.3 所示的小波包生成图中，位于同一级次上的波形具有相同的振荡频率，级次越高，振荡频率越高。

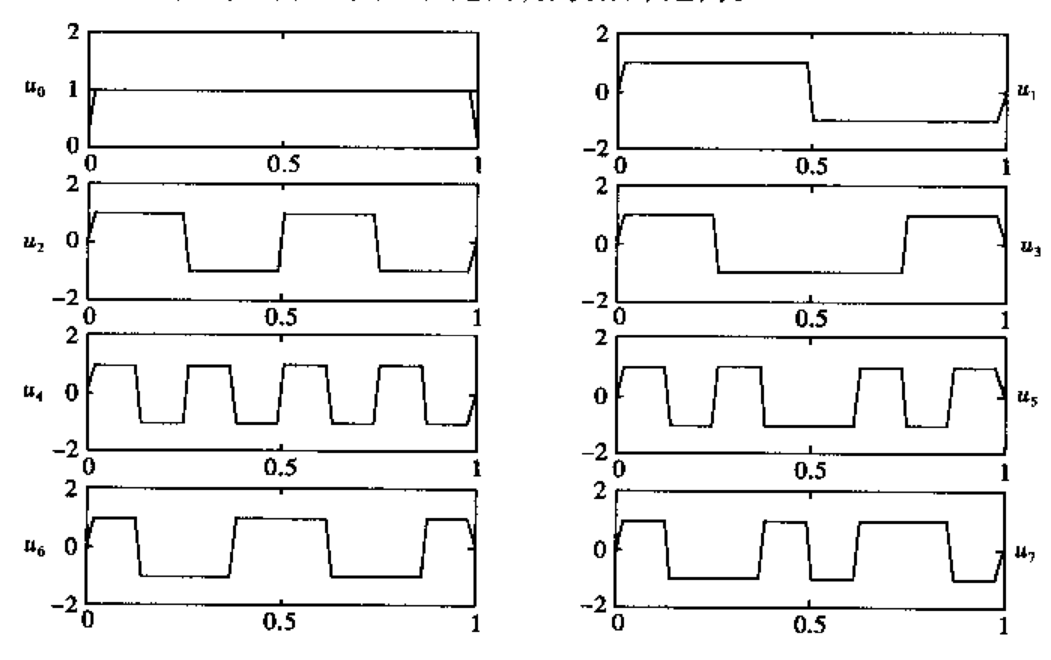


图 6.4　Harr 小波包波形

6.2　小波包分析

6.2.1　小波包分析

在信号的多分辨率分析中，信号空间的分解关系为

$$V_j = V_{j+1} \oplus W_{j+1} \tag{6.2.1}$$

它与(6.1.1)式和(6.1.2)式所表示的二尺度方程相对应。

如令

$$U_j^0 = V_j, U_j^1 = W_j; j \in Z \tag{6.2.2}$$

则(6.2.1)式可改写为

$$U_j^0 = U_{j+1}^0 \oplus U_{j+1}^1 \tag{6.2.3}$$

这实际上是把尺度函数空间和小波空间统一起来了。正如前面将二尺度方程(6.1.1)式和(6.1.2)式推广为(6.1.4)式和(6.1.5)式一样,也可以将(6.2.3)式推广为如下形式:

$$U_j^n = U_{j+1}^{2n} \oplus U_{j+1}^{2n+1}, \qquad n = 0,1,2,\cdots \tag{6.2.4}$$

它与(6.1.4)式和(6.1.5)式所表示的二尺度方程相对应。在(6.2.4)式中,当 $n=0$ 时,就退化为多分辨率分析。

在信号的多分辨率分析中,将尺度函数 ϕ 和正交小波 ψ 进行二进伸缩和平移之后形成 V_j 和 W_j 的正交基。类似的,也可以将 u_n 进行二进伸缩和平移得到

$$u_{j,n,k}(t) = 2^{-\frac{j}{2}} u_n(2^{-j}t - k) \tag{6.2.5}$$

定义 6.2 小波库的定义:称函数族 $\{u_{j,n,k}(t)\}_{j,k\in Z;n\in Z^+}$ 为小波库。

也就是说,小波库是将小波包中每个函数都进行二进伸缩和平移之后得到的函数族。

下述定理将证明:由正交镜像滤波器 h 和 g 生成的小波库也具有正交性。对于二进伸缩和平移,设 $\theta_j(t) = 2^{-\frac{j}{2}}\theta(2^{-j}t)$,则 $\theta_j(t-2^jk) = 2^{-\frac{j}{2}}\theta(2^{-j}t-k)$。

定理 6.4 设 $\{\theta_j(t-2^jk)\}_{k\in Z}$ 是空间 U_j 的正交规范基。由正交镜像滤波器 h 和 g 定义

$$\theta_{j+1}^0(t) = \sum_k h(k)\theta_j(t-2^jk) \tag{6.2.6}$$

$$\theta_{j+1}^1(t) = \sum_k g(k)\theta_j(t-2^jk) \tag{6.2.7}$$

则函数族

$$\{\theta_{j+1}^0(t-2^{j+1}k), \theta_{j+1}^1(t-2^{j+1}k), k \in Z\} \tag{6.2.8}$$

也是空间 U_j 的正交规范基。

证明 首先证明 $\{\theta_j(t-2^jk)\}_{k\in Z}$ 正交规范的频域表达形式是

$$\frac{1}{2^j}\sum_{l=-\infty}^{+\infty}\left|\hat{\theta}_j\left(\frac{\omega+2\pi l}{2^j}\right)\right|^2 = 1 \tag{6.2.9}$$

因为 $\{\theta_j(t-2^jk)\}_{k\in Z}$ 是正交规范基,所以

$$\begin{aligned}\langle\theta_j(t), \theta_j(t-2^jk)\rangle &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}|\hat{\theta}_j(\omega)|^2 e^{i2^jk\omega}d\omega \\ &= \frac{1}{2\pi}\sum_{l=-\infty}^{+\infty}\int_{2\pi l}^{2\pi(l+1)}\frac{1}{2^j}\left|\hat{\theta}_j\left(\frac{\omega}{2^j}\right)\right|^2 e^{ik\omega}d\omega \\ &= \frac{1}{2\pi}\int_0^{2\pi}e^{ik\omega}\frac{1}{2^j}\sum_{l=-\infty}^{+\infty}\left|\hat{\theta}_j\left(\frac{\omega+2\pi l}{2^j}\right)\right|^2 d\omega = \delta_k\end{aligned}$$

从而可得(6.2.9)式。

对(6.2.6)式和(6.2.7)式两边取傅立叶变换,得

$$\hat{\theta}_{j+1}^0(\omega) = H(2^j\omega)\hat{\theta}_j(\omega) \tag{6.2.10}$$

$$\hat{\theta}_{j+1}^1(\omega) = G(2^j\omega)\hat{\theta}_j(\omega) \tag{6.2.11}$$

于是,由(6.2.10)式有

$$\begin{aligned}&\frac{1}{2^{j+1}}\sum_{l=-\infty}^{+\infty}\left|\hat{\theta}_{j+1}^0\left(\frac{\omega+2\pi l}{2^{j+1}}\right)\right|^2 \\ &= \frac{1}{2^{j+1}}\sum_{l=-\infty}^{+\infty}\left|H\left(\frac{\omega}{2}+\pi l\right)\right|^2\left|\hat{\theta}_j\left(\frac{\omega/2+\pi l}{2^j}\right)\right|^2\end{aligned}$$

$$= \frac{1}{2}\left|H\left(\frac{\omega}{2}\right)\right|^2 \sum_{p=-\infty}^{+\infty} \frac{1}{2^j}\left|\hat{\theta}_j\left(\frac{\omega/2 + 2\pi p}{2^j}\right)\right|^2$$

$$+ \frac{1}{2}\left|H\left(\frac{\omega}{2} + \pi\right)\right|^2 \sum_{p=-\infty}^{+\infty} \frac{1}{2^j}\left|\hat{\theta}_j\left(\frac{\omega/2 + \pi + 2\pi p}{2^j}\right)\right|^2$$

$$= \frac{1}{2}\left|H\left(\frac{\omega}{2}\right)\right|^2 + \frac{1}{2}\left|H\left(\frac{\omega}{2} + \pi\right)\right|^2 = 1$$

这就是说$\{\theta_{j+1}^0(t-2^{j+1}k)\}_{k\in Z}$是正交规范基。同理，也可以证明$\{\theta_{j+1}^1(t-2^{j+1}k)\}_{k\in Z}$是正交规范基。上面用到了 h,g 的正交性：

$$|H(\omega)|^2 + |H(\omega+\pi)|^2 = 2 \qquad 或 \qquad |G(\omega)|^2 + |G(\omega+\pi)|^2 = 2$$

仿照上述过程，可以证明：如$\{\theta_{j+1}^0(t-2^{j+1}k)\}_{k\in Z}$与$\{\theta_{j+1}^1(t-2^{j+1}l)\}_{l\in Z}$正交，即

$$\langle \theta_{j+1}^0(t-2^{j+1}k), \theta_{j+1}^1(t-2^{j+1}l)\rangle = 0 \tag{6.2.12}$$

意味着

$$\frac{1}{2^{j+1}}\sum_{l=-\infty}^{+\infty} \hat{\theta}_{j+1}^0\left(\frac{\omega+2\pi l}{2^{j+1}}\right)\hat{\theta}_{j+1}^{1\,*}\left(\frac{\omega+2\pi l}{2^{j+1}}\right) = 0 \tag{6.2.13}$$

将(6.2.10)式和(6.2.11)式代入上式左边，并利用 h,g 之间的正交性

$$H(\omega)G^*(\omega) + H(\omega+\pi)G^*(\omega+\pi) = 0$$

可以证明(6.2.13)式成立，所以$\{\theta_{j+1}^0(t-2^{j+1}k)\}_{k\in Z}$与$\{\theta_{j+1}^1(t-2^{j+1}l)\}_{l\in Z}$正交。

为证明函数族$\{\theta_{j+1}^0(t-2^{j+1}k), \theta_{j+1}^1(t-2^{j+1}k)\}_{k\in Z}$与$\{\theta_j(t-2^jk)\}_{k\in Z}$生成相同的空间，必须证明：对任何 $a(k)$，总存在 $b(k)$ 和 $c(k)$，使得

$$\sum_{k=-\infty}^{+\infty} a(k)\theta_j(t-2^jk) = \sum_{k=-\infty}^{+\infty} b(k)\theta_{j+1}^0(t-2^{j+1}k) + \sum_{k=-\infty}^{+\infty} c(k)\theta_{j+1}^1(t-2^{j+1}k) \tag{6.2.14}$$

为此，对上式两边取傅立叶变换，得

$$A(2^j\omega)\hat{\theta}_j(\omega) = B(2^{j+1}\omega)\hat{\theta}_{j+1}^0(\omega) + C(2^{j+1}\omega)\hat{\theta}_{j+1}^1(\omega) \tag{6.2.15}$$

将(6.2.10)式和(6.2.11)式代入上式，并利用 $G(\omega) = e^{-i\omega}H^*(\omega+\pi)$，可以验证：若

$$B(2^{j+1}\omega) = \frac{1}{2}\left[A(2^j\omega)H^*(2^j\omega) + A(2^j\omega+\pi)H^*(2^j\omega+\pi)\right] \tag{6.2.16}$$

$$C(2^{j+1}\omega) = \frac{1}{2}\left[A(2^j\omega)G^*(2^j\omega) + A(2^j\omega+\pi)G^*(2^j\omega+\pi)\right] \tag{6.2.17}$$

则(6.2.15)式成立。证毕。

令 $\phi=\theta$，则定理6.4证明了多分辨率分解的空间分解关系式(6.2.1)式；如令 $u_n=\theta$，则定理6.4证明了小波包分解的空间分解关系式(6.2.4)式，而且还指出$\{u_{j,n,k}(t)\}_{k\in Z}$是 U_j^n 的正交规范基。其中，j 表示尺度为 2^j，n 为小波包参数，k 为平移参数。用 $u_{j,n,k}(t)$ 可以对信号在 $t=2^jk$ 附近的振动情况按尺度 2^j 进行分析。

在图6.2所式的小波包分析二叉树中，圆括号内的第一个参数为尺度（分辨率）参数 j，第二个参数为小波包（频率）参数 n。在每一节点处，函数族$\{u_{j,n,k}(t)\}_{k\in Z}$构成 U_j^n 的正交规范基。

小波包分析与多分辨率分析的不同之处就在于：不仅要对尺度函数空间 V_j 进行分解，同时也要对小波函数空间 W_j 进行分解。在图6.2所示的小波包二叉树分解结构中，圆括弧内第二个参数为1对应着小波函数空间 W_j，如果不继续分解下去，就是多分辨率分析，继续分解下

去就是小波包分析。

要实现对小波函数空间 W_L 的分解，将(6.2.4)式反复迭代就得到

$$\begin{aligned} W_L &= U_L^1 = U_{L+1}^2 \oplus U_{L+1}^3 \\ &= U_{L+2}^4 \oplus U_{L+2}^5 \oplus U_{L+2}^6 \oplus U_{L+2}^7 \\ &\cdots \quad \cdots \quad \cdots \\ &= U_{L+j}^{2^j} \oplus U_{L+j}^{2^j+1} \oplus \cdots \oplus U_{L+j}^{2^{j+1}-1} \\ &\cdots \quad \cdots \quad \cdots \\ &= U_{L+J}^{2^J} \oplus U_{L+J}^{2^J+1} \oplus \cdots \oplus U_{L+J}^{2^{J+1}-1} \end{aligned} \tag{6.2.18}$$

这种 W_L 空间分解的任意子空间序列表达式为

$$U_{L+j}^{2^j+p};j = 1,2,\cdots,J; \quad p = 0,1,\cdots,2^j - 1 \tag{6.2.19}$$

在对信号进行小波包分析时，若在尺度 2^L 逼近信号，则二叉树的树根为尺度空间 V_L。反复使用(6.2.4)式就得到小波包分析时的空间分解表达式

$$\begin{aligned} V_L &= U_L^0 = U_{L+1}^0 \oplus U_{L+1}^1 \\ &= U_{L+2}^0 \oplus U_{L+2}^1 \oplus U_{L+2}^2 \oplus U_{L+2}^3 \\ &\cdots \quad \cdots \quad \cdots \\ &= U_{L+j}^6 \oplus U_{L+j}^1 \oplus \cdots \oplus U_{l+j}^{2^j-1} \\ &\cdots \quad \cdots \quad \cdots \\ &= U_{L+J}^0 \oplus U_{L+J}^1 \oplus \cdots \oplus U_{L+J}^{2^J-1} \end{aligned} \tag{6.2.20}$$

这种 V_L 空间分解的任意子空间序列表达式为

$$U_{L+j}^n;j=0,1,2,\cdots,J; \quad n=0,1,\cdots,2^j-1 \tag{6.2.21}$$

其中，$L+j$ 表示尺度为 2^{L+j}，n 为小波包参数，j 表示分解的级次，所以 j 也表示节点在二叉树中的深度。与各节点对应的空间 U_{L+j}^n 的正交基为 $\{2^{-\frac{L+j}{2}}u_n(2^{-(L+j)}t-k)\}_{k\in Z}$。

如果树根与 V_0 相联系，那么，空间分解表达式为

$$\begin{aligned} V_0 &= U_0^0 = U_1^4 \oplus U_1^1 \\ &= U_2^0 \oplus U_2^1 \oplus U_2^2 \oplus U_2^3 \\ &\cdots \quad \cdots \quad \cdots \\ &= U_j^0 \oplus U_j^1 \oplus \cdots \oplus U_j^{2^j-1} \\ &\cdots \quad \cdots \quad \cdots \\ &= U_J^4 \oplus U_J^1 \oplus \cdots \oplus U_J^{2^J-1} \end{aligned} \tag{6.2.22}$$

这时 V_0 空间分解的任意子空间序列表达式为

$$U_j^n;j = 0,1,\cdots,J; \quad n = 0,1,\cdots,2^j - 1 \tag{6.2.23}$$

与之对应的正交基为

$$\{2^{-j/2}u_n(2^{-j}t - k);j = 0,1,\cdots,J;n = 0,1,\cdots;2^J - 1;k \in Z\} \tag{6.2.24}$$

上列两式中，j 既表示尺度为 2^j，又表示节点在二叉树中的深度，n 为小波包参数。图 6.2 为 3 级小波包分解的二叉树，如二叉树中的任何节点记为(j,n)，与各节点对应的正交基为 $\{2^{-j/2}u_n(2^{-j}t-k)\}_{k\in Z}$。

如 h 和 g 是支集长度为 N 的有限冲击响应正交镜像滤波器，那么，ϕ 和 ψ 的支集长度为$(N-$

1)；而且，所有 u_n 的支集长度均为$(N-1)$。所以，正交基中每个函数的支集长度均为$2^j(N-1)$。

也就是说，小波包分解的时域局部化与小波包参数无关。下面的例子说明了小波包分解的频域局部化。

例 6.2　Shannon 小波包。

Shannon 小波包由理想低通滤波器 h 和理想高通滤波器 g 生成，它们的傅立叶变换分别为

$$H(\omega)=\begin{cases}\sqrt{2}, & \omega\in[-\pi/2+2k\pi,\pi/2+2k\pi]\\ 0, & \text{其他}\end{cases}$$

和

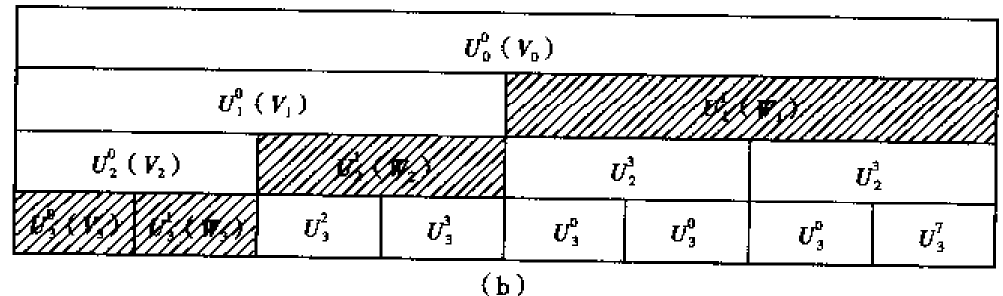
$$G(\omega)=\begin{cases}\sqrt{2}, & \omega\in[\pi/2+2k\pi,3\pi/2+2k\pi]\\ 0, & \text{其他}\end{cases}$$

尺度函数 ϕ 的傅立叶变换是

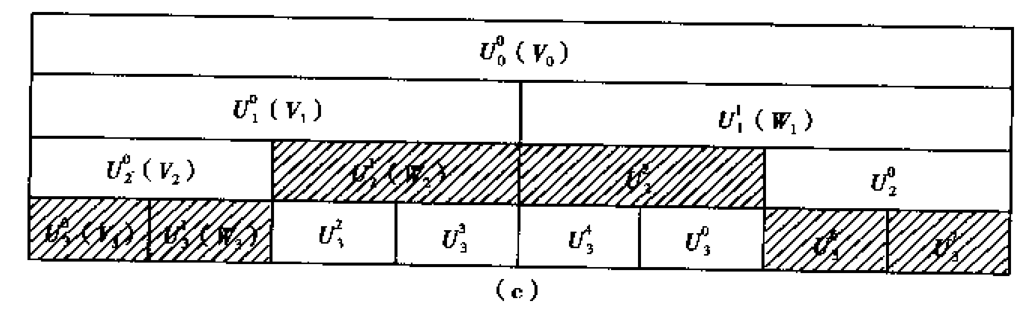
$$\hat{\phi}(\omega)=\begin{cases}1, & \omega\in[-\pi,\pi]\\ 0, & \text{其他}\end{cases}$$

$U_0^0(V_0)$							
$U_1^0(V_1)$				$U_1^1(W_1)$			
$U_2^0(V_2)$		$U_2^1(W_2)$		U_2^2		U_2^3	
$U_3^0(V_3)$	$U_3^1(W_3)$	U_3^2	U_3^3	U_3^4	U_3^5	U_3^6	U_3^7

(a)



(b)



(c)

图 6.5　(a) V_0 的子带基分解，(b) V_0 的多分辨率分解，(c) V_0 的任意小波包分解

由定理 6.4，在逐级分解的过程中，每次用 $H(2^j\omega)$ 和 $G(2^j\omega)$ 相乘实际上是将频率支集划分为两部分，图 6.5(a)表示在正频率$[0,\pi]$区间被逐级细分的过程，称为子带基分解。粗略地说，第 1 级将 $V_0=U_0^0$ 分解为 U_1^0 和 U_1^1，它们各占 V_0 频带的 1/2；第 2 级将 U_1^0 和 U_1^1 分解为 U_2^0，U_2^1

和 U_2^2, U_2^3，它们各占 V_0 频带的1/4；依次类推，第 j 级将 V_0 分解为 2^j 个子空间 $\{U_j^n; n=0,1,\cdots,2^j-1\}$，它们各占 V_0 频带的 $1/2^j$。更仔细地分析应该看到 $H(2^j\omega)$ 不总是起低通滤波的作用，这是由于伸缩，旁瓣被引入区间 $[0,\pi]$。

对其他类型的小波包，频域的划分与 Shannon 小波包类似。只不过 h 和 g 不是理想低通滤波器和理想高通滤波器，相邻频带的划分多少有些重叠而已。而且，由于旁瓣的作用，频域的划分要比时域的划分复杂得多，不能简单地认为在同一尺度级次上，按小波包参数的大小从左到右依次排列的节点所对应的小波包函数的中心频率是从小到大依次排列的。为了得到频率从小到大的依次排列，应将小波包参数转换为另一个节点参数，较详细的讨论可参阅文献[3]。另外，频域的划分显然与时间平移参数无关，时间平移参数只是将窗口平移以覆盖整个时间轴。

子带基分解并不是小波包分解的惟一选择，多分辨率分解的频率划分如图6.5(b)阴影区间所示，这些区间的并集正好覆盖了 V_0 的整个频带 $[0,\pi]$，可见多分辨率分解只不过是小波包分解的一种特殊选择方式。也可以按6.5(c)所示方式进行分解，阴影区间的并集也正好覆盖了 V_0 的整个频带。

6.2.2 小波包基

小波包分解实际上是在时间-频率平面的频率轴上逐级一分为二的过程，而对信号不同时间区间的分析是通过基函数的平移实现的，频率轴的划分在时间轴上是一致平移的。

如果要达到提高频域局部化的目的，应在 V_0 的频带内处处细分，这实际上是选择第 j 级子空间的基的集合来作为 V_0 的一组更精细的基

$$\{2^{-j/2}u_n(2^{-j}t-k); n=0,1,\cdots,2^j-1; k\in Z\} \tag{6.2.25}$$

如图6.5(a)所示，其中，$j=3$。但这并不是惟一的选择，例如多分辨率分析就是另一种可行的选择，如图6.5(b)所示。实际上，还有刚好覆盖 V_0 整个频带的其他选择方式，如图6.5(c)所示。这些选择既没丢失信息，又不存在信息冗余。

小波包分析时，基的选择的多种可能性也可从可容许树结构来理解。可容许树的每一个节点要么有零个子节点，要么有两个子节点。定理6.4指出可容许树中没有子节点的所有节点的基的并集是根节点的基。图6.6为一可容许树，它与图6.5(c)相对应，根节点 U_0^0 的基是 $\{U_2^1, U_2^2, U_3^0, U_3^1, U_3^0, U_3^7\}$ 的基的并集。

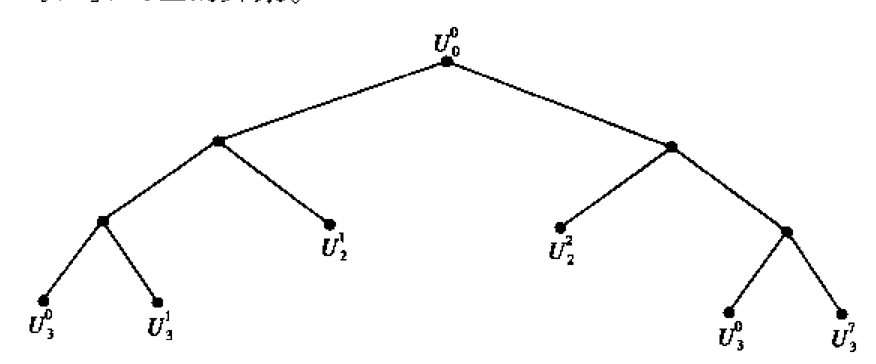


图6.6 小波包分析可容许树

由可容许树的定义，很容易证明如下定理，它给出了可容许树的充分必要条件。

定理6.5 设 $U_j^n, (j,n)\in S\subseteq Z\times Z^+$ 是 V_0 分解的子空间，即

$$V_0 = \bigoplus_{(j,n)\in S} U_j^n \tag{6.2.26}$$

使上式成立的充分必要条件是$[n/2^j,(n+1)/2^j),(j,n)\in S$是区间$[0,1)$的一个划分,即

$$\bigcup_{(j,n)\in S}[n/2^j,(n+1)/2^j)=[0,1) \tag{6.2.27}$$

且这些子区间两两不相交。

那么,到底有多少种可能的可容许树呢?或者说,小波包基到底有多少种可能的选择呢?下面的定理回答了这个问题。

定理6.6 假若小波包分解的深度为J,那么可容许树的数目B_J满足

$$2^{2^{J-1}}\leqslant B_J\leqslant 2^{\frac{5}{4}2^{J-1}} \tag{6.2.28}$$

对长度为N的离散信号,分解深度最多为$J=\log_2 N$,所以

$$2^{N/2}\leqslant B_J\leqslant 2^{\frac{5}{8}N} \tag{6.2.29}$$

对$L^2(R)$也同样存在多种频率轴的划分方式,而选择了某一种方式就意味着选择了相应的基,下面给出小波包基的定义。

定义6.3 小波包基:从小波包库$\{u_{j,n,k}(t)\}_{j,k\in Z;n\in Z^+}$中选取的能构成$L^2(R)$的一组正交规范基称为$L^2(R)$的一个小波包基。

当固定$n=1$时,$\{u_{j,n,k}(t)\}_{j,k\in Z;n\in 1}$实际上就是$\{\psi_{j,k}(t)\}_{j,k\in Z}$,它正是$L^2(R)$的正交小波基。所以,小波包基是小波基的推广形式。

由于

$$\begin{aligned}L^2(R)&=\cdots\oplus W_{L-2}\oplus W_{L-1}\oplus W_L\oplus W_{L+1}\oplus W_{L+2}\oplus\cdots\\&=\cdots\oplus W_{L-2}\oplus W_{L-1}\oplus W_L\oplus V_L\\&=\cdots\oplus U_{L-2}^1\oplus U_{L-1}^1\oplus U_L^1\oplus U_L^0\end{aligned}$$

由(6.2.4)式,有

$$\begin{aligned}U_{L-1}^1&=U_L^2\oplus U_L^3\\U_{L-2}^1&=U_{L-1}^2\oplus U_{L-1}^3=U_L^4\oplus U_L^5\oplus U_L^6\oplus U_L^7\\&\cdots\quad\cdots\quad\cdots\quad\cdots\end{aligned}$$

所以,可得

$$L^2(R)=\bigoplus_n U_L^n,n=0,1,2,\cdots \tag{6.2.30}$$

也就是说,对于某一固定的尺度2^L,$\{u_{j,n,k}(t);j=L,k\in Z,n=0,1,2,\cdots\}$是$L^2(R)$的一个小波包基,称为固定尺度的小波包基,或子带基。由于尺度固定,时-频窗是固定的,所以用固定尺度的小波包基分析信号类似于窗口傅立叶变换。

对$L^2(R)$小波包基的充分必要条件,有如下定理。

定理6.7 对于$S\subseteq Z\times Z^+$,$\{2^{-j/2}u_n(2^{-j}t-k)\}_{(j,n)\in S,k\in Z}$构成$L^2(R)$的小波包基的充分必要条件是$[n/2^j,(n+1)/2^j),(j,n)\in S$是区间$[0,1)$的一个划分。

6.2.3 小波包分析算法

对于正交小波包分析,其算法非常类似于正交小波分析的Mallat算法。

由于$\{u_{j,n,k}(t)\}_{k\in Z}$是空间$U_j^n$的正交规范基,所以信号$f(t)$在$U_j^n$的投影$f_{j,n}(t)$可表示为

$$f_{j,n}(t)=\sum_k d_{j,n}(k)u_{j,n,k}(t) \tag{6.2.31}$$

$$d_{j,n}(k)=\langle f(t),u_{j,n,k}(t)\rangle \tag{6.2.32}$$

称 $d_{j,n}(k)$ 为小波包系数。

$d_{0,0}$							
$d_{1,0}$				$d_{1,1}$			
$d_{2,0}$		$d_{2,1}$		$d_{2,2}$		$d_{2,3}$	
$d_{3,0}$	$d_{3,1}$	$d_{3,2}$	$d_{3,3}$	$d_{3,4}$	$d_{3,5}$	$d_{3,6}$	$d_{3,7}$

图 6.7 小波包分析 Mallat 算法树形图

因为 U_j^n 可以分解为 U_{j+1}^{2n} 和 U_{j+1}^{2n+1} 的直和，$U_j^n = U_{j+1}^{2n} \oplus U_{j+1}^{2n+1}$，且 U_{j+1}^{2n} 和 U_{j+1}^{2n+1} 的正交规范基分别是 $\{u_{j+1,2n,k}(t)\}_{k\in Z}$ 和 $\{u_{j+1,2n+1,k}(t)\}_{k\in Z}$。由二尺度方程(6.1.4)式和(6.1.5)式可得

$$u_{j+1,2n,k}(t) = \sum_l h(l-2k)u_{j,n,l}(t)$$

$$u_{j+1,2n+1,k}(t) = \sum_l g(l-2k)u_{j,n,l}(t)$$

令 $\bar{h}(l)=h(-l)$，$\bar{g}(l)=g(-l)$，并将上列两式两边与 f 求内积，有

$$d_{j+1,2n}(k) = \sum_l \bar{h}(2k-l)d_{j,n}(l) \tag{6.2.33}$$

$$d_{j+1,2n+1}(k) = \sum_l \bar{g}(2k-l)d_{j,n}(l) \tag{6.2.34}$$

它们就是小波包分解 Mallat 算法的表达式，如图 6.7 所示。

类似的，还可导出小波包重构算法的表达式如下

$$d_{j,n}(k) = \sum_l h(k-2l)d_{j+1,2n}(l) + \sum_l g(k-2l)d_{j+1,2n+1}(l) \tag{6.2.35}$$

6.3 最优基选择

既然小波包基有多种可能的选择，那么到底应该如何来选择小波包基呢？选择了一种小波包基，实际上就是选择了一种信号的描述方式，当然希望选择一种最有效的描述方式。小波变换是一种有效的方法，因为它将信号的能量集中在少数幅度较大的小波系数上。如选择一个适当的阈值，将小于阈值的小波系数置为0，只保留那些大于或等于阈值的小波系数用来逼近原始信号，逼近误差比较小。这种逼近与信号的性质有关，是一种非线性逼近。所以，“效率”高意味着非线性逼近误差小。小波包基具有更大的选择范围，而且它在频率域内做了更精致的细分，具有更好的自适应性，所以效率更高，非线性逼近误差更小。

6.3.1 花费函数

8.1 节将比较详细地讨论花费函数的定义，并指出花费函数度量了非线性逼近的均方误差，使花费函数最小的小波包基就是效率最高的，称为最优小波包基。显然，花费函数应该具有非负性和可加性。

定义 6.4 设序列 $d=\{d_k\}$ 是信号 $f(t)$ 在某一基下的系数，若存在非负连续函数 μ，使得

$C(f) = \sum_k \mu(|d_k|)$，$\mu(0) = 0$，则称 C 为 $f(t)$ 的一个花费函数。

常用的花费函数主要包括以下几种：

(1)令

$$\mu(x) = \begin{cases} 1, & x > T \\ 0, & x \leqslant T \end{cases} \tag{6.3.1}$$

这时 $C(f)$ 就是系数集 d 中绝对值大于阈值 T 的系数的个数。

(2)Shannon 熵

$$C(f) = -\sum_k d_k^2 \log(d_k^2) \tag{6.3.2}$$

且约定 $0\log 0 = 0$。上式是非归一化的 Shannon 熵。如令

$$p_k = \frac{|d_k|^2}{\|d\|^2} \tag{6.3.3}$$

则归一化的 Shannon 熵定义为

$$C(f) = -\sum_k p_k \log p_k \tag{6.3.4}$$

(3)能量对数熵

$$C(f) = \sum_k \log(d_k^2)$$

且约定 $\log 0 = 0$。

(4)对 $p \geqslant 1$，定义

$$C(f) = \sum_k |d_k|^p = \|d\|_p^p \tag{6.3.5}$$

这时，$C(f)$ 反映的是系数集 d 的"长度"。

6.3.2　最优基

最优小波包基是使花费函数最小的基。

定义 6.5　设实序列 $d = \{d_l\}$ 是 $f(t)$ 在 V_L 中的投影系数，B 是从小波库中选取的 V_L 的一个正交小波包基，Bd 是 d 在基 B 下的系数，若

$$C(\tilde{B}d) = \min\{C(Bd), B \in \mathbf{B}\} \tag{6.3.6}$$

则称 $\tilde{B}$ 是 d 的最优基，其中 $\mathbf{B}$ 表示 V_L 的所有小波包基。

对两个正交空间的正交规范基 B^0 和 B^1，花费函数在如下意义下是可加的

$$C(f, B^0 \cup B^1) = C(f, B^0) + C(f, B^1) \tag{6.3.7}$$

在小波包二叉树中，每一节点对应一个空间 U_j^n，其对应的正交基记为 B_j^n。该节点被分解为位于子节点上的两个正交子空间

$$U_j^n = U_{j+1}^{2n} \oplus U_{j+1}^{2n+1}$$

U_j^n 的最优基 O_j^n 是 U_j^n 所有基中使花费最小的那一个。下述定理给出了沿着树枝自下而上寻找最优基的一种迭代算法[3]。

定理 6.8　设 C 是一个可加的花费函数，则

$$O_j^n = \begin{cases} O_{j+1}^{2n} \oplus O_{j+1}^{2n+1}, & 若\ C(f, O_{j+1}^{2n}) + C(f, O_{j+1}^{2n+1}) < C(f, B_j^n) \\ B_j^n, & 若\ C(f, O_{j+1}^{2n}) + C(f, O_{j+1}^{2n+1}) \geqslant C(f, B_j^n) \end{cases} \tag{6.3.8}$$

根据这个定理，寻找最优基的迭代算法如下：设 V_L 为树根，j 表示树的深度，最大深度为 $J\geqslant 0$。若 $J=0$，则不进行分解，所以此时最优基为 $V_L=U_L^0$ 的基。

<table>
<tr><td colspan="8">50</td></tr>
<tr><td colspan="4">20</td><td colspan="4">22</td></tr>
<tr><td colspan="2">11</td><td colspan="2">12</td><td colspan="2">13</td><td colspan="2">14</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td></tr>
</table>

(a)

<table>
<tr><td colspan="8">50</td></tr>
<tr><td colspan="4">20</td><td colspan="4">22</td></tr>
<tr><td colspan="2">11</td><td colspan="2">12</td><td colspan="2">13</td><td colspan="2">14</td></tr>
<tr><td>1*</td><td>2*</td><td>3*</td><td>4*</td><td>5*</td><td>6*</td><td>7*</td><td>8*</td></tr>
</table>

(b)

<table>
<tr><td colspan="8">32 (50)</td></tr>
<tr><td colspan="4">10 (20)</td><td colspan="4">22*</td></tr>
<tr><td colspan="2">3 (11)</td><td colspan="2">7 (12)</td><td colspan="2">11 (13)</td><td colspan="2">14*</td></tr>
<tr><td>1*</td><td>2*</td><td>3*</td><td>4*</td><td>5*</td><td>6*</td><td>7*</td><td>8*</td></tr>
</table>

(c)

<table>
<tr><td colspan="8">32 (50)</td></tr>
<tr><td colspan="4">10 (20)</td><td colspan="4">22*</td></tr>
<tr><td colspan="2">3 (11)</td><td colspan="2">7 (12)</td><td colspan="2">11 (13)</td><td colspan="2">14*</td></tr>
<tr><td>1*</td><td>2*</td><td>3*</td><td>4*</td><td>5*</td><td>6*</td><td>7*</td><td>8*</td></tr>
</table>

(d)

图 6.8 (a) V_0 的小波包二叉树及花费函数，(b) 标记所有最底层节点，(c) 标记所有花费函数最小的节点，(d) 选取最优基，

当 $J>0$ 时，按以下步骤进行：

(1) 在树的底部，每一个 U_{L+J}^n 不再分解成子空间，因此 U_{L+J}^n 的惟一可用的基就是最优基，即：$O_{L+J}^n=B_{L+J}^n$。

(2)$J \geqslant j \geqslant 1, n = 0, 1, \cdots, 2^{j-1}-1$,按定理6.8,逐级迭代计算如下

$$O_{L+j-1}^{n} = \begin{cases} O_{L+j}^{2n} \oplus O_{L+j}^{2n+1}, & \text{若 } C(f, O_{L+j}^{2n}) + C(f, O_{L+j}^{2n+1}) < C(f, B_{L+j-1}^{n}) \\ B_{L+j-1}^{n}, & \text{若 } C(f, O_{L+j}^{2n}) + C(f, O_{L+j}^{2n+1}) \geqslant C(f, B_{L+j-1}^{n}) \end{cases}$$

(3)将 V_L 逐级分解为正交子空间 O_j^n 的直和,直到分解式中只有子空间 U_j^n,而没有 O_j^n。

例6.3 图6.8(a)是 $J=3$ 的小波包二叉树,各节点的花费函数列写于所有节点处。

(1)在最深级次的所有节点处作标记(本例为星号),如图6.8(b)。

(2)求两个子节点花费函数之和,如小于父节点的花费函数,则将该和值同时列写在父节点,否则,在父节点作星号标记,这样一直作到树根节点,如图6.8(c)。

(3)从上到下选取有星号标记的节点,如父节点有标记,则不再往下选取,这样选取到最底层。选中节点所对应的子空间的直和就构成 V_0 的最优基正交分解,而这些子空间的基的并集就是最优基,如图6.8(d)阴影所示。

例6.4 图6.9(a)为原始信号,图6.9(b)为其幅度谱。图6.9(b)的横轴为归一化数字频率,1.0对应奈奎斯特频率 f_n(即1/2采样频率),可以明显地看到原始信号包含两个正弦分量,其频率分别为 $0.09f_n$ 和 $0.95f_n$。

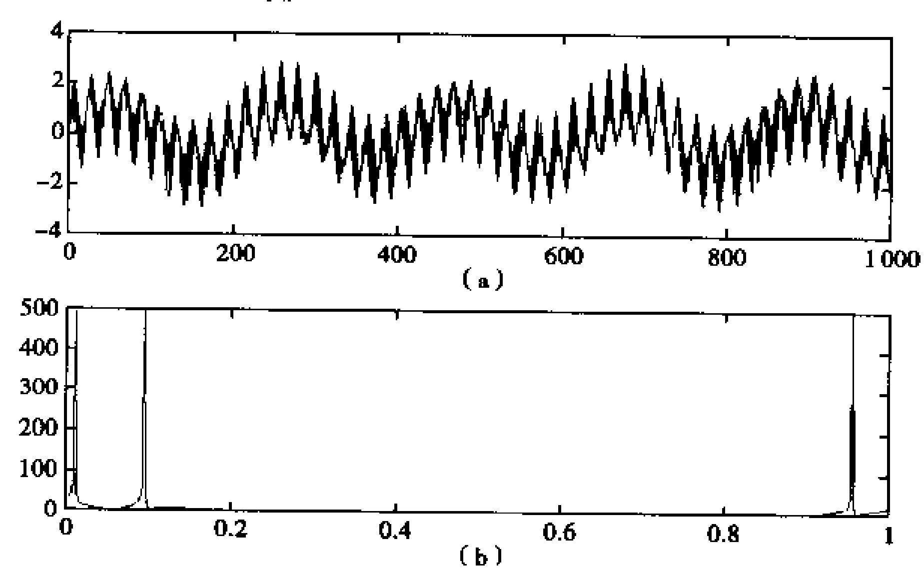


图6.9 (a)原始信号,(b)原始信号的频谱

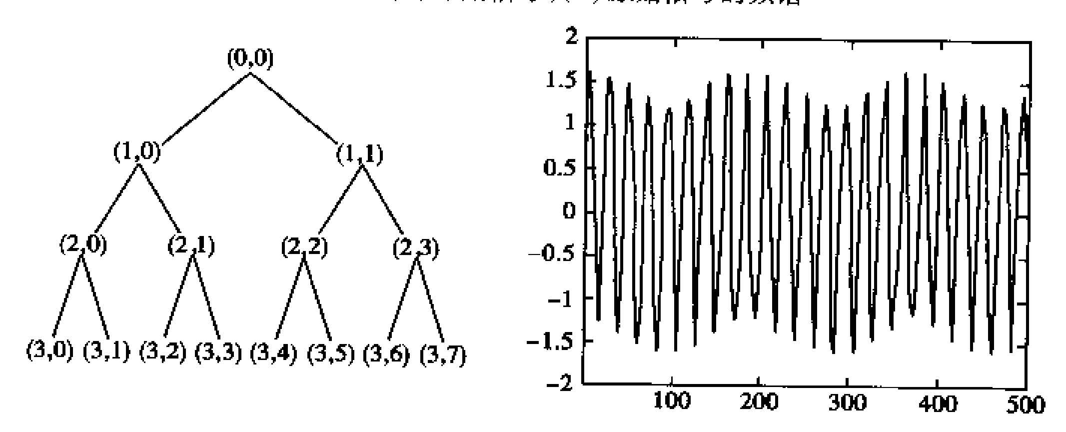


图6.10 小波包树及节点(1,1)的小波包系数

用Matlab小波工具箱中的函数wpdec对信号进行小波包分解,如图6.10所示。函数wp-

dec 返回小波包分解的树状结构，它是一个图形用户界面，在节点处点击即可得到相应节点系数的波形。原始信号有1 000个样本，节点(1.1)则有500个样本。可以用函数 wpcoef 得到树状结构中任意一个节点的系数值，图6.11画出了图6.10树状结构在节点(3,0)和(3,7)的系数值波形，它们的样本数均为125。

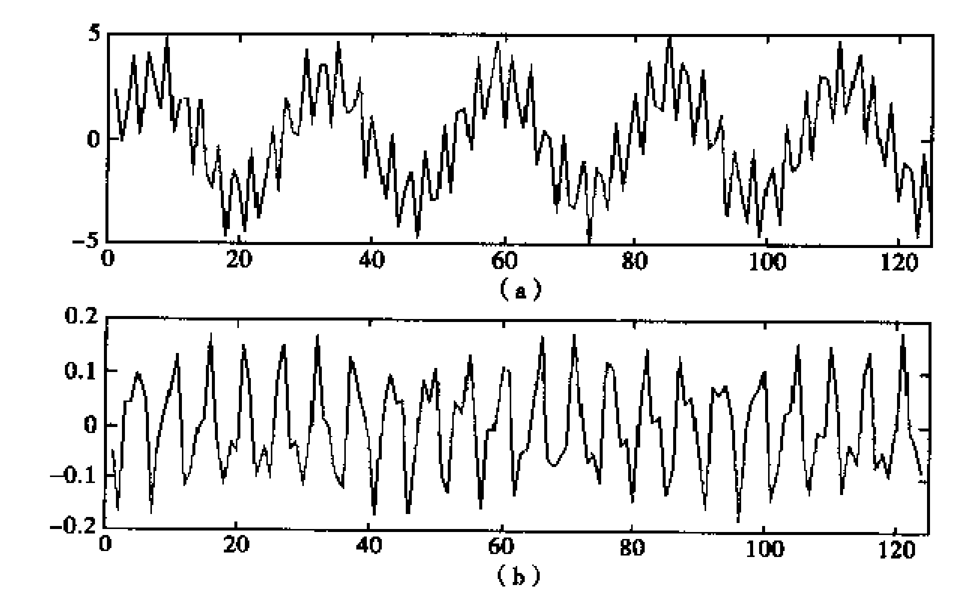


图6.11 (a)节点(3,0)的小波包系数，(b)节点(3,7)的小波包系数

6.4 图像小波包

6.4.1 小波包四叉树

图像多分辨率分析将原始图像分解成低频逼近图像和水平、垂直、斜角3个细节图像，再进一步逐级将低频逼近图像分解成4个图像。图像小波包分析除了分解低频逼近图像外，也要分解高频细节图像，这种分解过程表示为一个完整的四叉树，如图6.12所示。

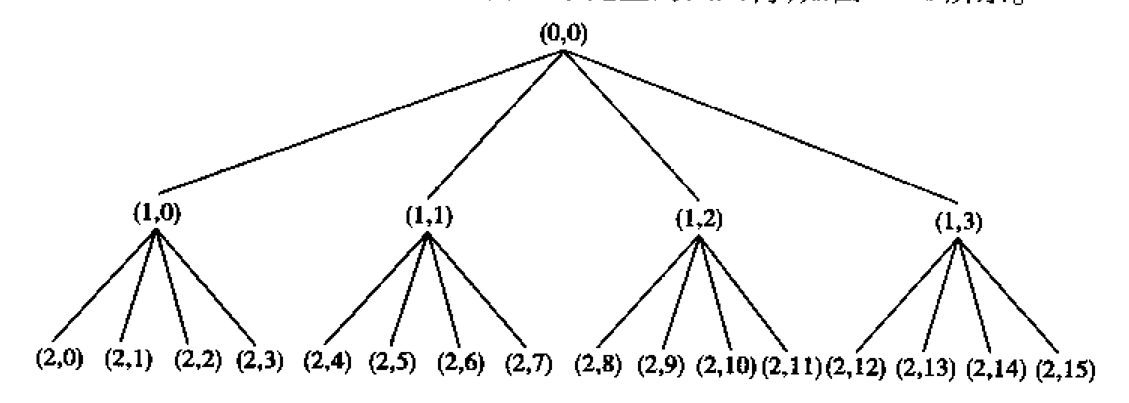


图6.12 图像小波包四叉树

四叉树的每个节点都对应一个可分离的二维函数空间，以 $U_j^{p,q}$ 表示之，

$$U_j^{p,q} = U_j^p \otimes U_j^q \tag{6.4.1}$$

其中，U_j^p 和 U_j^q 为一维小波包函数空间，j 为尺度参数，p 和 q 为小波包参数。$U_j^{p,q}$ 的正交规范基表示为

$$u_{j,n,m}^{p,q}(x,y) = u_{j,p,n}(x)u_{j,q,m}(y);\quad j,n,m \in Z;p,q \in Z^+ \tag{6.4.2}$$

其中 n 和 m 分别表示沿横向 x 和纵向 y 的空间平移参数。

一维小波包函数空间可分解成

$$U_j^p = U_{j+1}^{2p} \oplus U_{j+1}^{2p+1} \quad 和 \quad U_j^q = U_{j+1}^{2q} \oplus U_{j+1}^{2q+1} \tag{6.4.3}$$

代入(6.4.1)，有

$$U_j^{p,q} = U_{j+1}^{2p,2q} \oplus U_{j+1}^{2p,2q+1} \oplus U_{j+1}^{2p+1,2q} \oplus U_{j+1}^{2p+1,2q+1} \tag{6.4.4}$$

也就是说，一个二维函数空间分解成四个正交子空间的直和。

图6.12是Matlab小波工具箱图像小波包四叉树的表示方法，圆括号内的第1个参数表示尺度，第2个参数为小波包参数。

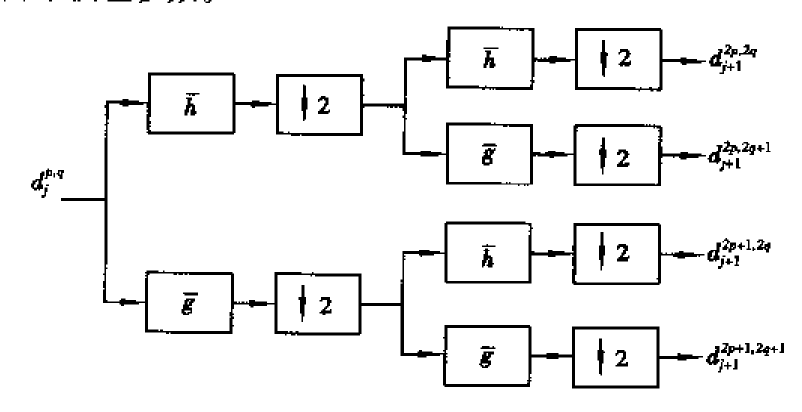


图6.13　图像小波包分解

6.4.2　图像小波包分析算法

从二维小波包基的可分离性及一维小波包分析算法可推出图像小波包分解算法

$$d_{j+1}^{2p,2q}(n,m) = d_j^{p,q} * \bar{h}\bar{h}(2n,2m) \tag{6.4.5}$$

$$d_{j+1}^{2p,2q+1}(n,m) = d_j^{p,q} * \bar{h}\bar{g}(2n,2m) \tag{6.4.6}$$

$$d_{j+1}^{2p+1,2q}(n,m) = d_j^{p,q} * \bar{g}\bar{h}(2n,2m) \tag{6.4.7}$$

$$d_{j+1}^{2p+1,2q+1}(n,m) = d_j^{p,q} * \bar{g}\bar{g}(2n,2m) \tag{6.4.8}$$

图像小波包分解算法如图6.13所示。

图像小波包重构算法为

$$\begin{aligned} d_j^{p,q}(n,m) = {} & \breve{d}_{j+1}^{2p,2q} * hh(n,m) + \breve{d}_{j+1}^{2p+1,2q} * hg(n,m) + \\ & \breve{d}_{j+1}^{2p,2q+1} * gh(n,m) + \breve{d}_{j+1}^{2p+1,2q+1} * gg(n,m) \end{aligned} \tag{6.4.9}$$

其中，$\breve{d}_{j+1}^{2p,2q}$ 表示对离散逼近图像 $d_{j+1}^{2p,2q}$ 按行和列插零，其余类推。图像小波包重构算法如图6.14所示。

6.4.3　最优图像小波包分析

在深度为 J 的完全小波包四叉树中，小波包基的个数 B_J 满足

$$2^{4^{J-1}} \leqslant B_J \leqslant 2^{\frac{49}{48}4^{J-1}} \tag{6.4.10}$$

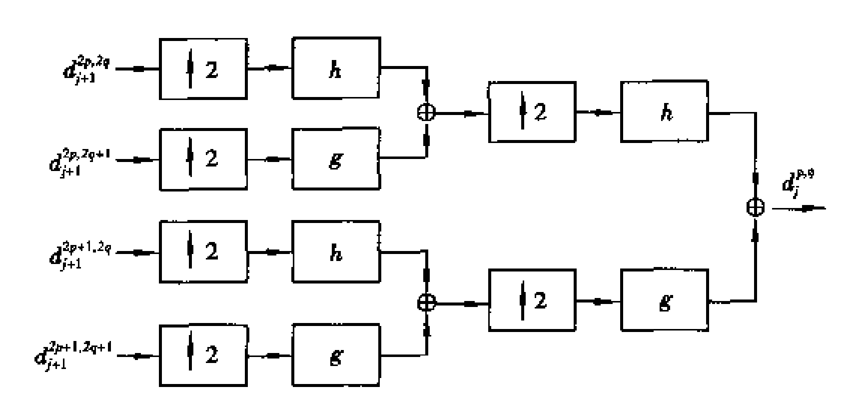


图 6.14　图像小波包重构

对有 $N\times N$ 个像素的图像，小波包四叉树的深度最多为 $\log_2 N$，因此小波包基的个数满足

$$2^{\frac{N^2}{4}}\leqslant B_j\log_2 N\leqslant 2^{\frac{49}{48}\frac{N^2}{4}} \tag{6.4.11}$$

与一维情况类似，定义花费函数来度量不同小波包基的效率，信号能量聚集在少量小波包系数上的程度越高，则效率越高，因为这些小波包系数足以充分刻画信号的内在性质。在用同样数量的系数逼近信号时，效率越高的小波包基，逼近的误差越小。关于好基的比较，有与定理 6.8 类似的定理。四叉树中任一节点对应的函数空间为 $U_j^{p,q}$，它的基为 $B_j^{p,q}$。对信号 f，用 $O_j^{p,q}$ 表示 $B_j^{p,q}$ 的最优基。由(6.4.4)式，$U_j^{p,q}$ 可分解成 4 个子空间，即 $B_j^{p,q}$ 可分解成 4 个子基的并集。

定理 6.9　设 C 是一个可加的花费函数，如果

$$C(f,B_j^{p,q}) < C(f,O_{j+1}^{2p,2q}) + C(f,O_{j+1}^{2p,2q+1}) + C(f,O_{j+1}^{2p+1,2q}) + C(f,O_{j+1}^{2p+1,2q+1})$$

则

$$O_j^{p,q} = B_j^{p,q}$$

否则

$$O_j^{p,q} = O_{j+1}^{2p,2q} \cup O_{j+1}^{2p,2q+1} \cup O_{j+1}^{2p+1,2q} \cup O_{j+1}^{2p+1,2q+1}$$

也就是说，在图像小波包分析四叉树结构中，如父节点的花费函数小于 4 个子节点花费函数之和，那么父节点的基是比 4 个子节点基的并集更好的基；否则，4 个子节点基的并集是比父节点基更好的基。按照这一定理，也可沿着树枝自下而上寻找最优基。

第 7 章 小波奇异性检测

信号中不规则的突变部分往往带有十分重要的信息，它是信号重要的特征之一。譬如，在故障诊断（特别是机械故障诊断）中，故障通常表现为输出信号发生突变，因而对突变点的检测在故障诊断中有着非常重要的意义。对图像信号而言，剧烈变化的部分如边缘和轮廓携带了重要的特征信息，所以对图像信号中边缘和轮廓的检测在图像数据压缩，模式识别和机器视觉等领域都有着非常重要的意义。长期以来，傅立叶变换是研究函数奇异性的主要工具，其方法是研究函数在傅立叶变换域的衰减以推断函数是否具有奇异性及奇异性的大小。但傅立叶变换缺乏空间局部性，它只能确定一个函数奇异性的整体性质，而难以确定奇异点在空间的位置及分布情况。由于小波变换具有空间局部性，它能“聚焦”于信号的局部结构，因此，利用小波变换来确定信号的奇异性位置更加有效。Lipschitz 指数被用来定量描述函数的奇异性。当尺度越来越精细时，小波变换模的极大值将收敛于信号的奇异点，其衰减速度取决于信号在奇异点的 Lipschitz 指数。所以小波变换不仅可以确定奇异发生的时间，而且可以进一步判断奇异的性质。

7.1 奇异性检测的基本原理

通常情况下，信号奇异性分两种情况：一种是信号在某一时刻，其幅值发生突变，引起信号的非连续，幅值突变处称为第一种类型的间断点；另一种是信号外观上很光滑，幅值没有突变，但是信号的一阶导数有突变发生，一阶导数不连续，称为第二种类型的间断点。

所谓尺度实际上是计算信号变化的范围。在图像处理中，边缘和轮廓的检测都是和一定的尺度相联系的。例如通过卫星图片检测陆地和海洋的分界线，尺度就比较大；而检测一个大型水库时尺度就比较小。也就是说，奇异性检测都是和一定的尺度相联系的，这与人的视觉特性是一致的。由于小波变换固有的“变焦”特性，所以非常适宜于图像的多尺度边缘检测。本节讨论一维信号的奇异性检测，然后将结果直接推广到二维情况。

7.1.1 Lipschitz 指数与奇异性

定义 4.1 说明 Lipschitz 指数可以来度量函数的正则性，如函数 f 在 t_0 的邻域内是一致

Lipschitz $\alpha > m$ 的，则 f 在该邻域内必然是 m 次连续可微的。也就是说，如 $\alpha > 1$，那么 f 至少是可微的。

如一个函数 f 在 t_0 点不可微，则说它在 t_0 点是奇异的。现在将 Lipschitz 指数引申到 $0 \leqslant \alpha < 1$，用以度量函数的奇异性。如函数 f 在 t_0 的邻域内是 $m = \lfloor \alpha \rfloor$ 次连续可微的，则(4.2.6)式中的多项式就是 f 在 t_0 的 $m-1$ 次泰勒展开式。当 f 在 t_0 奇异时，多项式 $P(t) = f(t)$，于是定义4.1引伸为如下定义。

定义 7.1 令 $0 \leqslant \alpha < 1$，如存在一个常数 C，使

$$\forall t \in R, |f(t) - f(t_0)| \leqslant C|t - t_0|^{\alpha} \tag{7.1.1}$$

成立，则称 f 在点 t_0 是 Lipschitzα 的。如对所有的 $t_0 \in [a,b]$ 和一个与 t_0 无关的常数 C，使得(7.1.1)成立，则称 f 在区间 $[a,b]$ 是一致 Lipschitzα 的。α 的上界值称为 Lipschitz 奇异性。

不难理解，如 f 在点 t_0 可微，则其 Lipschitz 指数至少为1。粗略地说，如 $\alpha = 1$，则(7.1.1)式可改写为 $\left|\dfrac{f(t) - f(t_0)}{t - t_0}\right| \leqslant C$，当 t 趋近于 t_0 时，不等式的左面实际上就是 f 在 t_0 点的一阶导数 $f'(t_0)$，取 $C \geqslant |f'(t_0)|$，则(7.1.1)式成立。

如 f 在点 t_0 不连续但在 t_0 的邻域有界，或者说它在 t_0 有有限跃度，则其 Lipschitz 指数为0。当 $\alpha = 0$ 时，(7.1.1)式成为 $|f(t) - f(t_0)| \leqslant C$，左面最多等于 f 在 t_0 点的跃度，取 C 等于或大于跃度，则(7.1.1)式成立。

还可以将 Lipschitz 指数推广到为负数的情况：如 f 的原函数在 t_0 点的 Lipschitz 指数为 α，则 f 在该点的 Lipschitz 指数为 $\alpha - 1$。例如 $\delta(t - t_0)$ 的原函数为一单位阶跃，它在 t_0 的 Lipschitz 指数为0，故 $\delta(t - t_0)$ 在 t_0 点的 Lipschitz 指数为 -1。现在可以清楚地看到，Lipschitz 指数确实能在更一般的意义下定量地描述函数的奇异性。

7.1.2 多尺度微分算子

为了能用小波变换准确检测信号奇异点的位置，希望小波有小的支集，即希望小波在时域快速衰减，这意味着对任意衰减指数 $m \in N$，存在 C_m，使得

$$\forall t \in R, |\psi(t)| \leqslant \frac{C_m}{1 + |t|^m} \tag{7.1.2}$$

在1.5节曾指出，若小波有 K 阶消失矩，则它与直到 $K-1$ 次的多项式正交，如函数 f 在 t_0 点足够正则，那么小波变换将取决于函数的 K 阶导数。下述定理明确指出，小波变换确实相当于一个微分算子。

定理 7.1[3] 快速衰减的小波 ψ 具有 K 阶消失矩，当且仅当存在快速衰减的函数 θ，使得

$$\psi(t) = (-1)^K \frac{\mathrm{d}^K}{\mathrm{d}t^K}\theta(t) \tag{7.1.3}$$

从而

$$Wf(s,t) = s^K \frac{\mathrm{d}^K}{\mathrm{d}t^K}(f * \bar{\theta}_s)(t) \tag{7.1.4}$$

其中，$\bar{\theta}_s(t) = s^{-1/2}\theta(-t/s)$。而且 ψ 具有不超过 K 阶的消失矩当且仅当 $\int_{-\infty}^{+\infty}\theta(t)\mathrm{d}t \neq 0$。

(7.1.4)式意味着小波变换等于用 $\bar{\theta}_s(t)$ 平滑信号后求 K 阶导数，所以它相当于一个多尺度微分算子。式中的小波变换为通常使用的相关型小波变换，而且本章都假定小波是实的，即

$$Wf(s,t) = \frac{1}{\sqrt{s}}\int_{-\infty}^{+\infty} f(\tau)\psi\left(\frac{\tau - t}{s}\right)\mathrm{d}\tau = f * \bar{\psi}_s(t) \tag{7.1.5}$$

其中，$s \in R^+$ 为尺度参数，$t \in R$ 为平移参数，$\bar{\psi}_s(\tau) = s^{-1/2}\psi(-\tau/s)$。

由于 θ 的积分不等于零，所以它是一个平滑函数。例如高斯函数就是一个经常用到的平滑函数，其数学表达式为

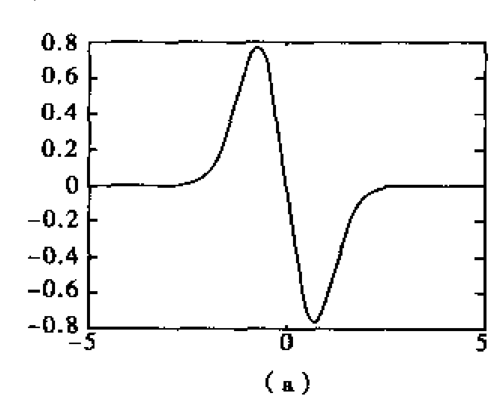
$$\theta(t) = \frac{1}{\sqrt{\pi}}\mathrm{e}^{-t^2}$$

其中，常数 $1/\sqrt{\pi}$ 是使 $\int_{-\infty}^{+\infty}\theta(t)\mathrm{d}t = 1$。高斯函数的 1 阶导数和 2 阶导数分别为

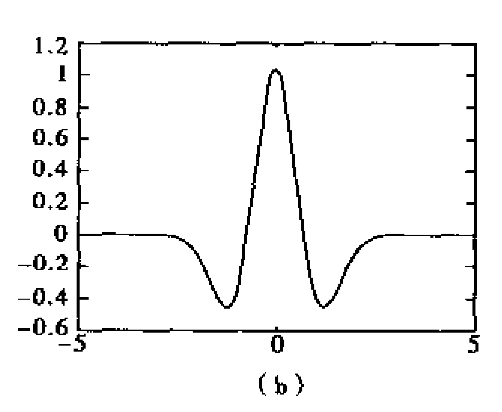
$$\psi^1(t) = -2\left(\frac{2}{\pi}\right)^{\frac{1}{4}} t e^{-t^2}$$

$$\psi^2(t) = -\frac{2}{\sqrt{3}}\left(\frac{2}{\pi}\right)^{\frac{1}{4}}(1 - 2t^2)e^{-t^2}$$

上列两式中的常数项是使其范数等于 1，图 7.1 为 ψ^1 和 ψ^2 的波形，它们是在信号奇异性检测中经常使用的小波。



(a)



(b)

图 7.1　(a)ψ^1，(b)ψ^2

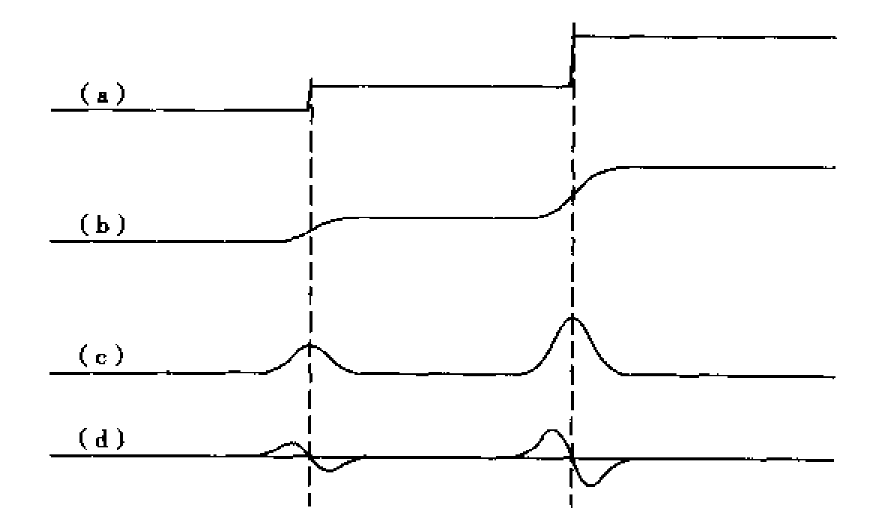


图 7.2　(a)$f(t)$，(b)$f * \bar{\theta}_s(t)$，(c)$W_1 f(s,t)$，(d)$W_2 f(s,t)$

大多数多尺度边缘检测器都是在不同的尺度平滑信号，然后由其1阶或2阶导数检测锐变点。所以边缘检测可以通过小波变换来实现。在平滑函数为高斯函数的情况下，2阶导数过零点检测等价于 Marr 的边缘检测法，而1阶导数极值点检测等价于 Canny 的边缘检测法 图7.2表示出用 ψ^1 和 ψ^2 检测边缘的过程，(a)为具有边缘的信号，它被高斯平滑后示于(b)，(c)和(d)分别为用 ψ^1 和 ψ^2 所做出的小波变换 $W_1f(s,t)$ 和 $W_2f(s,t)$ 的波形，$W_1f(s,t)$ 的极值点和 $W_2f(s,t)$ 的过零点检测出边缘的位置。$W_2f(s,t)$ 的过零点是 $f*\bar{\theta}_s(t)$ 的拐点。1阶导数极值点检测和2阶导数过零点检测虽然有相似之处，但前者却具有某些突出的优点。$f*\bar{\theta}_s(t)$ 的拐点可能是其1阶导数绝对值的极大值点，也可能是其1阶导数绝对值的极小值点。1阶导数绝对值的极大值点反应了 $f*\bar{\theta}_s(t)$ 的锐变，而极小值点却对应着较缓慢的变化。所以，Mallat 提出了寻找 $|Wf(s,t)|$ 的极大值点实现信号奇异性检测的方法。由图7.2可以看到，在尺度精细化的过程中，$|W_1f(s,t)|$ 和 $|W_2f(s,t)|$ 的极大值点都将收敛于信号的奇异点。下面将会看到，在尺度 $s\to 0$ 的过程中，$|Wf(s,t)|$ 的衰减速度取决于信号在奇异点的 Lipschitz 指数，故寻找1阶导数的极值点较寻找2阶导数过零点的方法会获得更多关于边缘的信息。

7.1.3 小波变换度量正则性

假定 θ 的面积为1，由于它具有快速衰减，所以

$$\lim_{s\to 0} s^{-1/2}\bar{\theta}_s=\delta$$

因此对任意一个在 t 点连续的函数 f，有

$$\lim_{s\to 0} f*(s^{-1/2}\bar{\theta}_s)=f$$

如果 f 在 t 的邻域内是 K 次连续可微的，则(7.1.4)式意味着

$$\lim_{s\to 0} s^{-(K+1/2)}Wf(s,t)=\lim_{s\to 0} f^{(K)}*s^{-1/2}\bar{\theta}_s(t)=f^{(K)}(t)$$

如 $f\in C^K$，则 $|Wf(s,t)|=O(s^{K+1/2})$。也就是说，小波变换在细尺度下的衰减速度取决于信号的正则性。当 f 在 t 的邻域内不可微，下述定理进一步将小波变换的衰减速度和 Lipschitz 指数联系起来，它即适合函数正则，也适合函数奇异的情况。

定理 7.2[3] 设 $f(t)\in L^2(R)$ 在区间 $[a,b]$ 是一致 Lipschitz $\alpha\leqslant K$ 的，总存在 $A>0$ 使得

$$\forall(s,t)\in R^+\times[a,b],\ |Wf(s,t)|\leqslant As^{\alpha+1/2} \tag{7.1.6}$$

反之，若 f 有界且对某一非整数的 $\alpha<K$，$Wf(s,t)$ 满足上式，则对 $\forall\varepsilon>0$，f 在 $[a+\varepsilon,b-\varepsilon]$ 上是一致 Lipschitz α 的。

(7.1.6)式中的小波变换是相关型小波变换。在 Mallat 早期关于奇异性检测的文章[6，7]中，用的是卷积型小波变换，这时(7.1.6)式的右边修改为 As^{α} 即可。

上述定理实际上是 $s\to 0$ 时，$|Wf(s,t)|$ 衰减速度与 Lipschitz 指数的关系，对大尺度，Cauchy-Schwarz 不等式指出小波变换是有界的

$$|Wf(s,t)|=|\langle f,\psi_{s,t}\rangle|\leqslant\|f\|\cdot\|\psi\|$$

这时小波变换的模将不受(7.1.6)式的约束。

上述定理是关于区间 Lipschitz 指数的，下述定理则将一点的 Lipschitz 指数与小波变换的衰减速度联系起来。

定理 7.3[3] 设 $f(t)\in L^2(R)$ 在 t_0 点是 Lipschitz $\alpha\leqslant K$ 的，总存在常数 $A>0$ 使得

$$\forall(s,t)\in R^+\times R,\ |Wf(s,t)|\leqslant As^{\alpha+1/2}\left(1+\left|\frac{t-t_0}{s}\right|^{\alpha}\right) \tag{7.1.7}$$

反之,设非整数的 $\alpha < K$,且存在 A 和 $\alpha' < \alpha$ 使得

$$\forall (s,t) \in R^{+} \times R, |Wf(s,t)| \leqslant A s^{\alpha+1/2}\left(1+\left|\frac{t-t_0}{s}\right|^{\alpha'}\right) \tag{7.1.8}$$

则 f 在点 t_0 是 Lipschitzα 的。

如 ψ 是紧支的,支集为 $[-c,c]$,则 $s^{-1/2}\psi(t/s)$ 的支集为 $[-sc,sc]$。在时间-尺度平面上,当时间平移量 t 满足

$$|t-t_0| \leqslant cs \tag{7.1.9}$$

时,分析小波的支集才包含 t_0 点。(7.1.9)式规定了时间-尺度平面上的一个锥形区域,称为点 t_0 的影响锥。在影响锥内,有 $|t-t_0|/s \leqslant c$,这时条件(7.1.7)式和(7.1.8)式可写成

$$|Wf(s,t)| \leqslant A' s^{\alpha+1/2} \tag{7.1.10}$$

因为幅值较大的小波系数落在奇异点的影响锥内,所以用小波变换度量 Lipschitz 指数时,定理7.3与定理7.2的一致 Lipschitz 条件完全相同。

上面两个定理指出,随着尺度越来越精细,小波变换模将呈指数衰减,Lipschitz 指数越大,衰减越快。这样,就可以由小波变换模的衰减速度来度量 Lipschitz 指数。我们知道,可以通过频率趋于无穷大时傅立叶变换的衰减速度来度量函数的正则性,这与小波变换是相通的,因为尺度越小意味着频率越高。但傅立叶变换只能判断函数的整体正则性,而小波变换不仅能度量函数在某一区间的局部正则性,而且还能度量函数在某一点的正则性。当用小波变换来检测信号的奇异性时,不仅可以判断奇异的性质(测出 Lipschitz 指数),而且还能确定发生奇异的时间。

例7.1 有一信号 $f(t)$ 在 t_0 点出现单位阶跃,如图7.3(a)所示。用 $-\psi^1$ 和 ψ^2 对它做小波变换得 $W_1 f(s,t)$ 和 $W_2 f(s,t)$,令 $-\psi_s^1(\tau) = -s^{-1/2}\psi^1(\tau/s)$,$\psi_s^2(\tau) = s^{-1/2}\psi^2(\tau/s)$,则小波变换可以解释为 $f(\tau)$ 与 $-\psi_s^1(\tau-t)$ 和 $\psi_s^2(\tau-t)$ 相乘之后的面积,即它们在 $\tau \geqslant t_0$ 部分的面积。如图7.3(b)所示,$W_1 f(s,t)$ 在 $t=t_0$ 点出现极大,极大值 $W_1 f(s,t_0)$ 为图中阴影部分面积,它等于 $-\psi^1$ 正半部分的面积乘以 $s^{1/2}$,由(7.1.6)式可以判断信号在 t_0 点的 Lipschitz 指数 $\alpha=0$。设 Δ 为 ψ^2 中心到过零点的长度,$W_2 f(s,t)$ 则在 $t=t_0 \pm s\Delta$ 处出现极小和极大。图7.3(c)表示 $t=t_0-s\Delta$ 的情况,这时 $W_2 f(s,t)$ 现极大值,它等于 ψ^2 中心到 Δ 部分的面积乘以 $s^{1/2}$,由此仍可判断信号在 t_0 点的 Lipschitz 指数 $\alpha=0$。当 $s \to 0$ 时,$|W_2 f(s,t)|$ 的极大值点都将收敛于信号

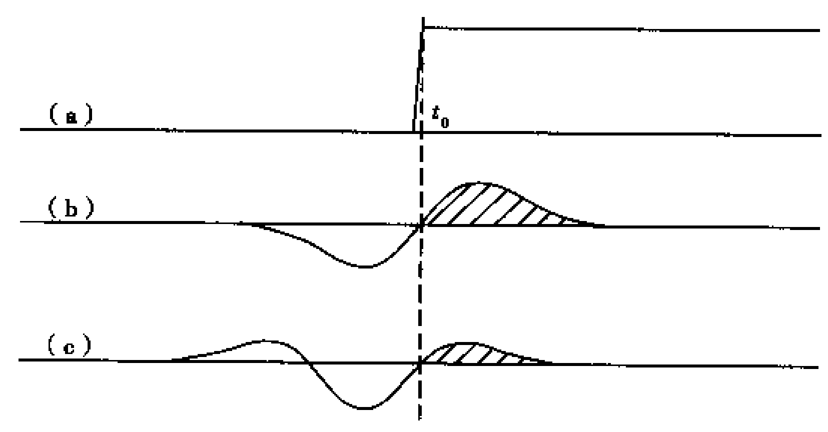


图7.3 (a)$f(\tau)$,(b) $-\psi_s^1(\tau-t_0)$,(c)$\psi_s^2(\tau-t_0+s\Delta)$

的奇异点。

如信号为$f(t)=\delta(t-t_0)$，则其小波变换 $Wf(s,t)=\bar{\psi}_s(t-t_0)$。用 $-\psi^1$ 做小波变换时，$|W_1f(s,t)|$会出现两个极大值，而用 ψ^2 做小波变换时，$|W_2f(s,t)|$会出现 3 个极大值，且它们的幅度均与 $s^{-1/2}$成正比，故信号在 t_0 点的 Lipschitz 指数 $\alpha=-1$。而且当 $s\to 0$ 时，小波变换模极大都将收敛于信号的奇异点。

7.2 奇异性检测

本节讨论信号孤立奇异性检测，它包含两点：寻找小波变换模极大在细尺度时收敛的坐标以确定奇异发生的时间；通过小波变换模极大的衰减计算 Lipschitz 指数，判断奇异的性质。小波变换模极大意味着小波变换对时间平移量的偏导数等于 0。

如小波具有 1 阶消失矩，那么小波变换模极大就是信号被平滑后的 1 阶导数的模极大，这时小波变换模极大可用来确定一维信号的不连续点（跳变点）或图像的边缘。那么，其他性质的奇异点如何呢？

定理 7.4[3]　设 ψ 是紧支的 C^K 函数，且 $\psi=(-1)^K\theta^{(K)}$，$\int_{-\infty}^{+\infty}\theta(t)\mathrm{d}t\neq 0$。设 $f\in L^1[a,b]$，如存在 $s_0>0$，以致 $t\in[a,b]$及 $s<s_0$ 时$|Wf(s,t)|$没有局部极大，则对任意 $\varepsilon>0$，函数f在$[a-\varepsilon,b+\varepsilon]$上是一致 Lipschitz K 的。

此定理意味着：如信号f在 t_0 点奇异（Lipschitz 指数小于 1），则必然在 t_0 的邻域内会出现小波变换模极大序列，它们在细尺度下收敛于 t_0。

在时间－尺度平面上将这些模极大值点连接起来得到模极大线，但在奇异点两侧可能会出现多条模极大线（与所使用的小波有关），在细尺度下这些模极大线收敛点的坐标就是奇异发生的时刻（或位置）。在收敛的过程中，模极大线有可能中断，也就是说在某些尺度时，奇异点附近没有模极大。可以证明[3]：当平滑函数 θ 是高斯函数时，模极大线总是连续的。此外，定理 7.4 也并不保证出现模极大线的点就一定是奇异点，实际情况也确实是这样。所以在出现模极大线的情况下，应进一步计算 Lipschitz 指数，从而判断该点是否奇异；如该点确实是奇异点，则可用 Lipschitz 指数衡量该点奇异的性质。

对(7.1.6)式两边取对数，得

$$\log_2|Wf(s,t)|\leqslant\log_2 A+(\alpha+1/2)\log_2 s \tag{7.2.1}$$

由上式可以看到，小波变换模的极大值是随尺度 s 而变化的，如按对数取值，$\log_2|Wf(s,t)|$与 $\log_2 s$ 具有线性关系，它们之间的比例系数（直线的斜率）即为 Lipschitz 指数。

例 7.2　用高斯函数的 1 阶导数做连续小波变换，检测信号的奇异性。

图 7.4(a)为原始信号。取尺度 $1\leqslant s\leqslant 32$ 做连续小波变换，图 7.4(b)为 $s=32$ 时的小波变换值，可以明显地看到出现了 4 个模极大值点。为了确定奇异点的位置，应找出模极大线，如图 7.4(c)所示。在 $n=245$ 和 $n=372$ 附近，模极大值点基本与尺度无关（由于计算误差，模极大值点存在一定扰动）。而在 $n=718$ 附近则出现了两条模极大线，当尺度变得越来越精细时，这两条模极大线越来越靠近 $n=718$，这是因为信号在该处近似于一个尖脉冲，又使用了具有双峰的小波，所以出现了两条模极大线。因此在利用小波变换作奇异性检测时，应根据对奇

异性的先验知识和经验选用适当的小波，对变化剧烈的奇异应选用支撑宽度窄的小波，而对变化较平缓的奇异可选用支撑宽度宽的小波。如对奇异性没有任何先验知识和经验，则应反复试验并进行理论分析以确定合适的小波。

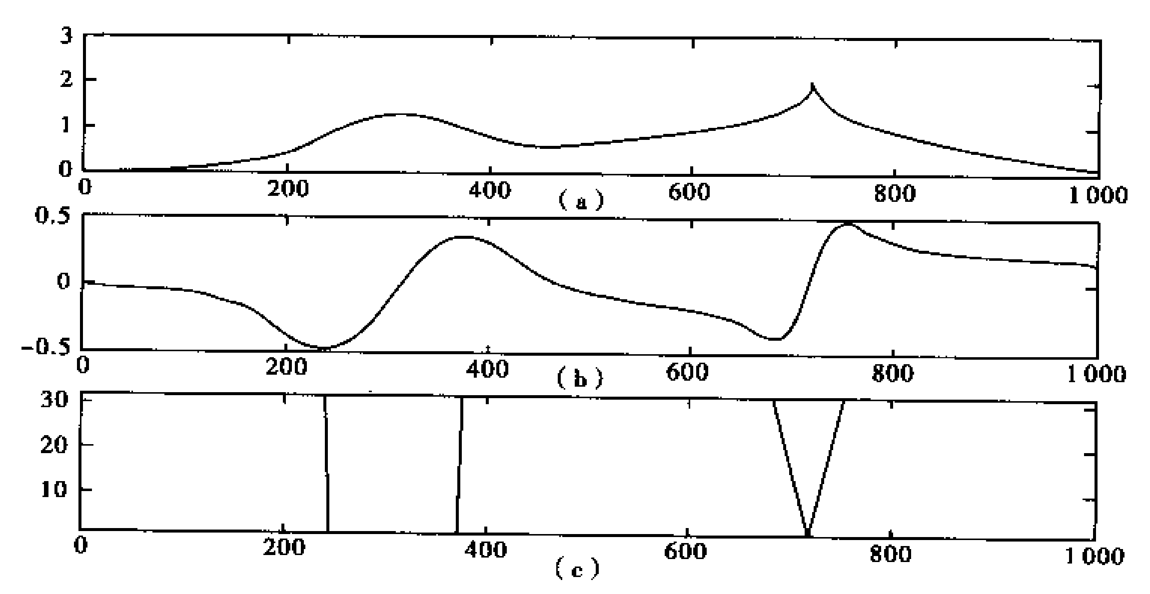


图 7.4　(a)信号，(b)小波变换，(c)模极大线

为了进一步确定奇异的性质，应作出模极大值与尺度的对数关系曲线，如图 7.5 所示。图中横坐标为尺度 $1 \leqslant s \leqslant 32$ 时的对数，纵坐标为小波变换模极大值的对数。正如(7.2.1)式所示，模极大值与尺度的对数关系曲线是一条直线，而直线的斜率就是描述奇异性的 Lipschitz 指数。图 7.5 中，曲线 1 为 $n=245$ 处的对数关系曲线，其 Lipschitz 指数 $\alpha=0.96$。曲线 2 为 $n=372$ 处的对数关系曲线，Lipschitz 指数 $\alpha=0.95$。所以 $n=245$ 和 $n=372$ 这两点不是奇异点。曲线 3 和 4 均为 $n=718$ 处的对数关系曲线，Lipschitz 指数 $\alpha=0.37$。我们知道，如函数在某一点 1 阶可微，则其 Lipschitz 指数至少等于 1；如在某一点不连续，则其 Lipschitz 指数等于 0。信号在 $n=718$ 处 Lipschitz 指数 $\alpha=0.37$，说明信号在该点不可微但连续。

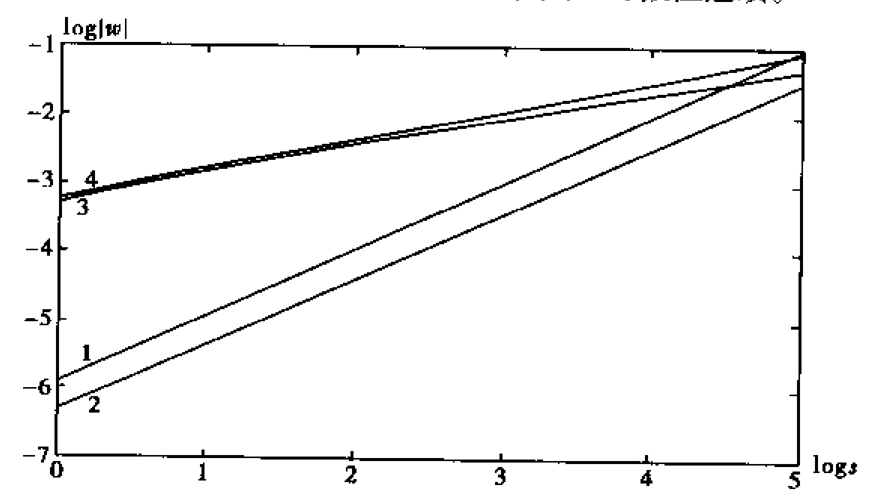


图 7.5　模极大值与尺度的对数关系曲

在实际应用中，并不存在纯数学意义下的阶跃、冲击这类奇异性，相反，信号中有些连续可微的地方却可能含有重要的信息。例如在图像的边界处灰度有很快的变化，但由于衍射作用，

它却不是不连续的。对于这类物理现象,可以建立如下数学模型。

假设函数$f(t)$在t_0有急剧变化,可以认为

$$f(t)=f_0(t)*g_\sigma(t) \tag{7.2.2}$$

其中,f_0在t_0有严格数学意义下的奇异性,它在t_0的邻域内是一致 Lipschitz α 的,g_σ是方差为σ^2的高斯函数,即

$$g_\sigma(t)=\frac{1}{\sqrt{2\pi}}\sigma\exp\left(\frac{-t^2}{2\sigma^2}\right)$$

如小波是高斯平滑函数

$$\theta(t)=\lambda\exp\left(\frac{-t^2}{2\beta^2}\right) \tag{7.2.3}$$

的导数,$\psi=(-1)^K\theta^{(K)}$,则可以证明[3]

$$|Wf(s,t)|\leqslant As^{\alpha+1/2}\left(1+\frac{\sigma^2}{\beta^2s^2}\right)^{-(K-\alpha)/2} \tag{7.2.4}$$

当尺度相对很大即$s\gg\sigma/\beta$时,小波变换的模仍按$s^{\alpha+1/2}$衰减,这时$\psi_s(\tau)=s^{-1/2}\psi(\tau/s)$的支集远大于$g_\sigma$的支集,从大范围看,信号仍是奇异的。而当尺度相对很小即$s\ll\sigma/\beta$时,信号却是无限可微的,小波变换的模按$s^{K+1/2}$衰减。对上式取对数,有

$$\log_2|Wf(s,t)|\leqslant\log_2A+(\alpha+1/2)\log_2s-\frac{(K-\alpha)}{2}\log_2(1+\sigma^2/\beta^2s^2) \tag{7.2.5}$$

7.3 离散二进小波变换的快速算法

上面讨论了用连续小波变换做奇异性检测。由于二进小波变换未对时间平移参数离散化,具有时间平移不变性,也适宜于做奇异性检测。而且离散二进小波变换可以通过 à Trous 算法快速实现,它类似于双正交小波变换中的 Mallat 算法。

7.3.1 二进小波设计

在2.4节曾指出,二进小波ψ及其对偶$\tilde{\psi}$必须满足

$$\sum_{j=-\infty}^{+\infty}\hat{\psi}^*(2^j\omega)\hat{\tilde{\psi}}(2^j\omega)=1 \tag{7.3.1}$$

和构造双正交小波类似,仍从滤波器组的设计着手。设h和g是一对有限冲击响应滤波器,且h是满足$H(0)=\sqrt{2}$的低通滤波器,构造傅立叶变换满足

$$\hat{\phi}(\omega)=\prod_{p=1}^{+\infty}\frac{H(2^{-p}\omega)}{\sqrt{2}}=\frac{1}{\sqrt{2}}H\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \tag{7.3.2}$$

的尺度函数。因为$H(0)=\sqrt{2}$,所以$\hat{\phi}(0)=1$。且假定上列无穷级数收敛于一个能量有限函数,从而$\phi\in L^2(R)$。相应的,小波ψ也由傅立叶变换的二尺度方程定义为

$$\hat{\psi}(\omega)=\frac{1}{\sqrt{2}}G\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \tag{7.3.3}$$

由于$\hat{\psi}(0)=0$,所以$G(0)=0$,g是一个高通滤波器。

对应的 $\tilde{\phi}$ 和重构小波 $\tilde{\psi}$ 也由重构滤波器 $\tilde{h}$ 和 $\tilde{g}$ 定义如下

$$\hat{\tilde{\phi}}(\omega)=\prod_{p=1}^{+\infty}\frac{\tilde{H}(2^{-p}\omega)}{\sqrt{2}}=\frac{1}{\sqrt{2}}\tilde{H}\left(\frac{\omega}{2}\right)\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right) \tag{7.3.4}$$

$$\hat{\tilde{\psi}}(\omega)=\frac{1}{\sqrt{2}}\tilde{G}\left(\frac{\omega}{2}\right)\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right) \tag{7.3.5}$$

其中，$\tilde{h}$ 是一个低通滤波器，$\tilde{g}$ 是一个高通滤波器。

需要强调的是，这儿的 ϕ 和 $\tilde{\phi}$ 虽然称为尺度函数，但并不一定要求它们生成多分辨分析。

为使 ψ 和 $\tilde{\psi}$ 满足(7.3.1)，滤波器组必须满足一定条件。下述定理给出一个充分条件。

定理 7.5　若滤波器组满足

$$\forall\omega\in[-\pi,\pi],\tilde{H}(\omega)H^*(\omega)+\tilde{G}(\omega)G^*(\omega)=2 \tag{7.3.6}$$

则

$$\forall\omega\in R-\{0\},\sum_{j=-\infty}^{+\infty}\hat{\psi}^*(2^j\omega)\hat{\tilde{\psi}}(2^j\omega)=1 \tag{7.3.7}$$

证明　由(7.3.3)式和(7.3.5)式可得

$$\hat{\tilde{\psi}}(\omega)\hat{\psi}^*(\omega)=\frac{1}{2}\tilde{G}\left(\frac{\omega}{2}\right)G^*\left(\frac{\omega}{2}\right)\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right)\hat{\phi}^*\left(\frac{\omega}{2}\right)$$

而由(7.3.6)式可得

$$\hat{\tilde{\psi}}(\omega)\hat{\psi}^*(\omega)=\frac{1}{2}\left[2-\tilde{H}\left(\frac{\omega}{2}\right)H^*\left(\frac{\omega}{2}\right)\right]\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right)\hat{\phi}^*\left(\frac{\omega}{2}\right)=\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right)\hat{\phi}^*\left(\frac{\omega}{2}\right)-\hat{\tilde{\phi}}(\omega)\hat{\phi}^*(\omega)$$

令 l,k 为正整数，由上式可得

$$\begin{aligned}\sum_{j=-l}^{+k}\hat{\tilde{\psi}}(2^j\omega)\hat{\psi}^*(2^j\omega)&=\sum_{j=-l}^{+k}\hat{\tilde{\phi}}(2^{j-1}\omega)\hat{\phi}^*(2^{j-1}\omega)-\sum_{j=-l}^{+k}\hat{\tilde{\phi}}(2^j\omega)\hat{\phi}^*(2^j\omega)\\&=\hat{\tilde{\phi}}(2^{-l-1}\omega)\hat{\phi}^*(2^{-l-1}\omega)-\hat{\tilde{\phi}}(2^k\omega)\hat{\phi}^*(2^k\omega)\end{aligned}$$

当 $\omega\neq0$ 时，令 $l,k\to+\infty$，则因 $\hat{\phi}(0)=\hat{\tilde{\phi}}(0)=1$，而 $\omega\to\infty$ 时 $\hat{\phi}(\omega)$ 和 $\hat{\tilde{\phi}}(\omega)$ 均趋近于 0，可由上式得(7.3.7)式。证毕。

上述定理中的(7.3.6)式与双正交小波理想重构滤波器的单位增益条件(5.2.11)式完全一样，但不要求去混叠条件(5.2.10)式，因为二进小波变换未在时域采样。不难理解，正交小波和双正交小波都是二进小波。为构造紧支二进小波，要求滤波器是有限冲击响应的，也可用 5.4 节类似的方法设计 h 和 $\tilde{h}$，但并不要求满足(5.2.23)式。

7.3.2　样条二进小波

与设计样条双正交小波类似，如取 ϕ 为 m 次样条函数，其傅立叶变换

$$\hat{\phi}(\omega)=e^{-\frac{i\varepsilon\omega}{2}}\left(\frac{\sin(\omega/2)}{(\omega/2)}\right)^{m+1} \tag{7.3.8}$$

那么

$$H(\omega)=\sqrt{2}e^{-i\varepsilon\omega/2}[\cos(\omega/2)]^{m+1} \tag{7.3.9}$$

当 m 为奇数时，$\varepsilon=0$，ϕ 以 $t=0$ 为对称中心，h 以 $n=0$ 为对称中心；当 m 为偶数时，$\varepsilon=1$，ϕ 以

$t=1/2$ 为对称中心，h 以 $n=1/2$ 为对称中心。

为构造有 1 阶消失矩的小波，应选择在 $\omega=0$ 的邻域内 $G(\omega)=O(\omega)$，例如

$$G(\omega) = -i\sqrt{2}e^{-i\omega/2}\sin(\omega/2) \tag{7.3.10}$$

从而有

$$\hat{\psi}(\omega) = \frac{-i\omega}{4}\exp\left(\frac{-i\omega(1+\varepsilon)}{4}\right)\left(\frac{\sin(\omega/4)}{(\omega/4)}\right)^{m+2} \tag{7.3.11}$$

二进样条小波可解释为：当 m 为偶数时，将 $m+1$ 次样条延迟 1，再压缩 1/2 后取 1 阶导数的负值得到 ψ，其中心为 1/2；当 m 为奇数时，直接将 $m+1$ 次样条压缩 1/2 后取 1 阶导数的负值得到 ψ，这时中心为 1/4。所以，ψ 的中心位于 $(1+\varepsilon)/4$，且关于中心反对称。图 7.6 画出了 $m=2$ 时的样条小波及相应的尺度函数。

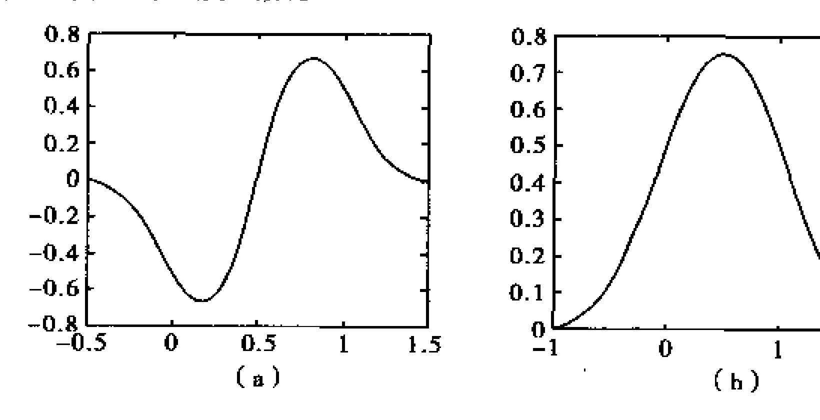


图 7.6　$m=2$(a)样条小波，(b)尺度函数

为设计对偶样条函数 $\hat{\phi}$ 和 $\hat{\psi}$，取 $\tilde{H}(\omega)=H(\omega)$，从而由(7.3.6)式得

$$\tilde{G}(\omega) = \frac{2-|H(\omega)|^2}{G^*(\omega)} = -i\sqrt{2}e^{-i\omega/2}\sin(\omega/2)\sum_{n=0}^{m}[\cos(\omega/2)]^{2n} \tag{7.3.12}$$

当 m 为奇数时，令 $m=2k-1$，则有

$$H(\omega) = \sqrt{2}[\cos(\omega/2)]^{2k} = \frac{\sqrt{2}}{4^k}\sum_{n=-k}^{+k}\binom{2k}{k-n}e^{-in\omega} \tag{7.3.13}$$

所以

$$h(n) = \frac{\sqrt{2}}{4^k}\binom{2k}{k-n}, \qquad -k \leqslant n \leqslant +k \tag{7.3.14}$$

当 m 为偶数时，令 $m=2k$，则有

$$H(\omega) = \sqrt{2}e^{-i\omega/2}[\cos(\omega/2)]^{2k+1} = \frac{\sqrt{2}}{2^{2k+1}}\sum_{n=-k}^{k+1}\binom{2k+1}{k-n+1}e^{-in\omega} \tag{7.3.15}$$

所以

$$h(n) = \frac{\sqrt{2}}{2^{2k+1}}\binom{2k+1}{k-n+1}, \qquad -k \leqslant n \leqslant k+1 \tag{7.3.16}$$

表 7.1 列出了 $m=2$ 时的滤波器组。

表 7.1　$m=2$ 时的二进样条滤波器组

n	$h/\sqrt{2}$	$\tilde{h}/\sqrt{2}$	$g/\sqrt{2}$	$\tilde{g}/\sqrt{2}$
-2				-0.031 25
-1	0.125	0.125	-0.5	-0.218 75
0	0.375	0.375	0.5	-0.687 5
1	0.375	0.375		0.687 5
2	0.125	0.125		0.218 75
3				0.031 25

7.3.3　à Trous 算法

相关型二进小波变换定义为

$$Wf(2^j,t) = 2^{-j/2}\int_{-\infty}^{+\infty} f(\tau)\psi(2^{-j}(\tau - t))\mathrm{d}\tau, \qquad j \in Z \tag{7.3.17}$$

令 $\bar{\psi}_j(\tau) = \psi_j(-\tau)$，$\psi_j(\tau) = 2^{-j/2}\psi(2^{-j}\tau)$，则相关型二进小波变换也可写成卷积形式

$$Wf(2^j,t) = f * \bar{\psi}_j(t), \qquad j \in Z \tag{7.3.18}$$

离散二进小波变换定义为对时间平移量采样，假定采样间隔为 1，即 $t=n$，则有

$$d_j(n) = Wf(2^j,n) = f * \bar{\psi}_j(t)\big|_{t=n} \tag{7.3.19}$$

这里的 $d_j(n)$ 形式上有些类似正交多分辨分析中的离散细节信号，但仔细比较可以看出，离散二进小波变换在任何尺度下都是同样的采样间隔。同样定义

$$a_j(n) = f * \bar{\phi}_j(t)\big|_{t=n} \tag{7.3.20}$$

其中，$\bar{\phi}_j(\tau) = \phi_j(-\tau)$，$\phi_j(\tau) = 2^{-j/2}\phi(2^{-j}\tau)$。

采样定理告诉我们，时域采样意味着频域重复后叠加，令 $\hat{a}_j(\omega)$ 和 $\hat{d}_j(\omega)$ 分别表示 $a_j(n)$ 和 $d_j(n)$ 的傅立叶变换，则由(7.3.20)式可得

$$\hat{a}_j(\omega) = \sum_k \hat{f}(\omega + 2k\pi)\hat{\phi}_j^*(\omega + 2k\pi) \tag{7.3.21}$$

其中，$\hat{\phi}_j^*(\omega) = 2^{j/2}\hat{\phi}^*(2^j\omega)$ 为 $\bar{\phi}_j(\tau)$ 的傅立叶变换。由(7.3.2)式有

$$2^{j/2}\hat{\phi}^*[2^j(\omega + 2k\pi)] = H^*[2^{j-1}(\omega + 2k\pi)]2^{(j-1)/2}\hat{\phi}^*[2^{j-1}(\omega + 2k\pi)]$$

当 $j \geqslant 1$ 时，$H[2^{j-1}(\omega + 2k\pi)] = H(2^{j-1}\omega)$，上式可写成

$$\hat{\phi}_j^*(\omega + 2k\pi) = H^*(2^{j-1}\omega)\hat{\phi}_{j-1}^*(\omega + 2k\pi)$$

将它代入(7.3.21)式可得

$$\hat{a}_j(\omega) = H^*(2^{j-1}\omega)\sum_k \hat{f}(\omega + 2k\pi)\hat{\phi}_{j-1}^*(\omega + 2k\pi) = H^*(2^{j-1}\omega)\hat{a}_{j-1}(\omega)$$

令 $H_j(\omega) = H(2^j\omega)$，则由上式得 $\hat{a}_{j+1}(\omega) = H_j^*(\omega)\hat{a}_j(\omega)$，其傅立叶反变换为

$$a_{j+1}(n) = \bar{h}_j * a_j(n) \tag{7.3.22}$$

类似的，令 $G_j(\omega) = G(2^j\omega)$，有

$$d_{j+1}(n) = \bar{g}_j * a_j(n) \tag{7.3.23}$$

上列两式便是离散二进小波变换的分解算法，称为 à trous 算法。其中，$\bar{h}_j(n) = h_j(-n)$，$\bar{g}_j(n) =$

$g_j(-n)$。由于 $H_j(\omega)$ 是 $h_j(n)$ 的傅立叶变换，故 $h_j(n)$ 是将 $h(n)$ 插入 2^j-1 个零点的结果，即

$$h_j(n) = \begin{cases} h(2^{-j}n), & n = 2^j k, k \in Z \\ 0, & \text{其他} \end{cases} \tag{7.3.24}$$

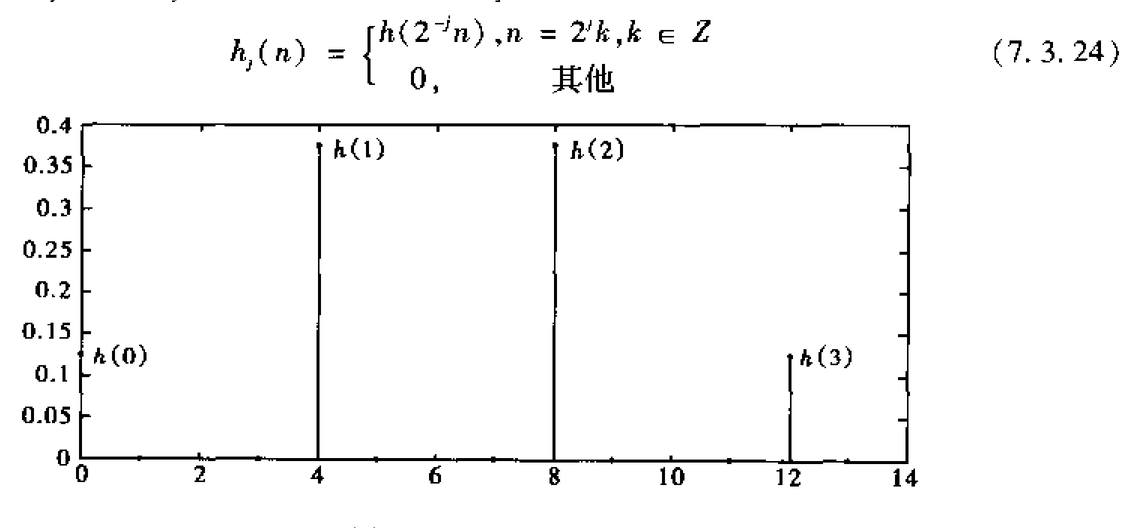


图 7.7 $h_j(n), j=2$

如图 7.7 所示。同理，$G_j(\omega)$ 是 $g_j(n)$ 的傅立叶变换，故

$$g_j(n) = \begin{cases} g(2^{-j}n), & n = 2^j k, k \in Z \\ 0, & \text{其他} \end{cases} \tag{7.3.25}$$

由条件(7.3.6)，有

$$\tilde{H}_j(\omega)H_j^*(\omega)\hat{a}_j(\omega) + \tilde{G}_j(\omega)G_j^*(\omega)\hat{a}_j(\omega) = 2\hat{a}_j(\omega)$$

而(7.3.22)式和(7.3.23)式意味着 $\hat{a}_{j+1}(\omega) = H_j^*(\omega)\hat{a}_j(\omega)$，$\hat{d}_{j+1}(\omega) = G_j^*(\omega)\hat{a}_j(\omega)$，将它们代入上式得

$$\hat{a}_j(\omega) = \frac{1}{2}[\tilde{H}_j(\omega)\hat{a}_{j+1}(\omega) + \tilde{G}_j(\omega)\hat{d}_{j+1}(\omega)]$$

取其傅立叶反变换得离散二进小波变换的重构算法

$$a_j(n) = \frac{1}{2}[\tilde{h}_j * a_{j+1}(n) + \tilde{g}_j * d_{j+1}(n)] \tag{7.3.26}$$

假定原始信号为 $a_0(n)$，则可以把它分解成

$$[\{d_j\}_{1\leq j\leq J}, a_J] \tag{7.3.27}$$

也能从它们重构 $a_0(n)$。$H_j(\omega)$ 和 $G_j(\omega)$ 是 $H(\omega)$ 和 $G(\omega)$ 向低频方向移动的结果，所以离散二进小波变换系数 $\{d_j\}_{1\leq j\leq J}$ 刻画了信号在不同尺度时的剧烈变化，而 a_J 则保留了大尺度下的低频信息。

离散二进小波变换快速算法与正交和双正交多分辨分析的 Mallat 算法确有相似之处，但区别也是很明显的，离散二进小波分解算法中不容许抽取偶数样本，否则会丢失信息，重构算法中也不必插零。此外，à trous 算法中的分解和重构滤波器随着尺度变化要膨胀和压缩。可以用正交和双正交滤波器组实现 à trous 算法，这时显然有信息冗余，但正因为如此，才保证了奇异点检测的准确性。

离散二进小波变换可理解为连续小波变换在离散网格 $s=2^j, t=n$ 上的采样，由(7.1.6)式得

$$|Wf(2^j, n)| \leq A(2^j)^{\alpha+1/2} \tag{7.3.28}$$

根据(7.3.19)式的定义，上式又可写成

$$|d_j(n)| \leqslant A(2^j)^{\alpha+1/2} \tag{7.3.29}$$

将上式两边取对数有

$$\log_2|d_j(n)| \leqslant \log_2 A + (\alpha + 1/2)j \tag{7.3.30}$$

这意味着离散二进小波变换模极大收敛于信号的奇异点，模极大值的对数与尺度 j 成正比，从比例系数可算出 Lipschitz 指数。

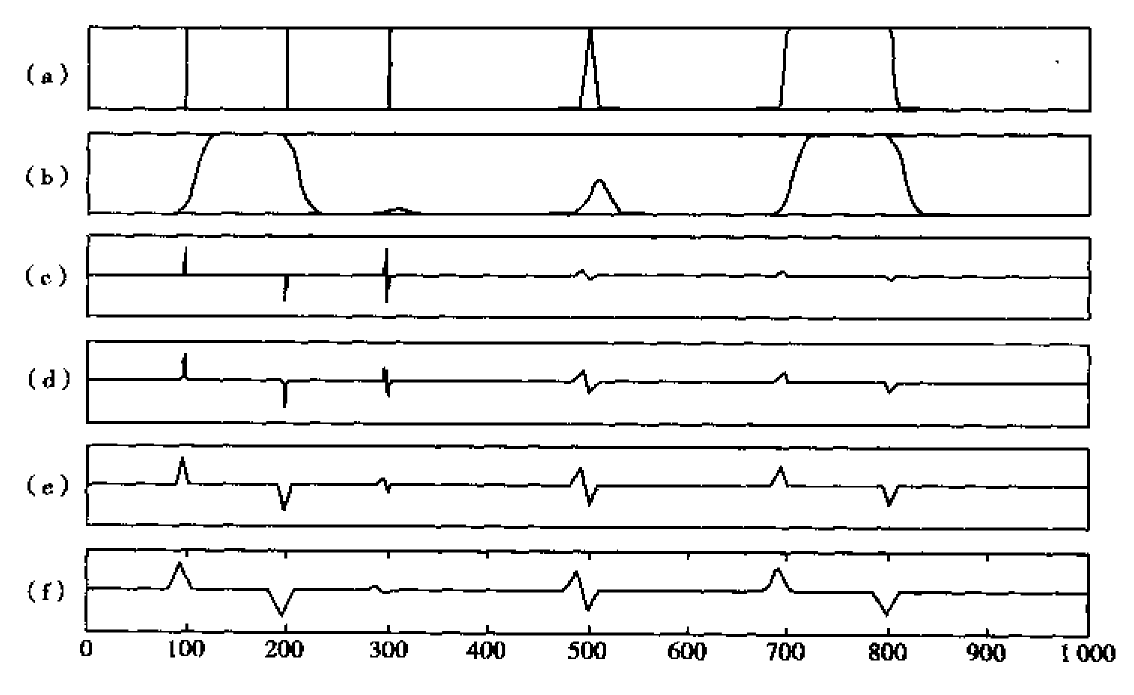


图 7.8　奇异性检测

(a) a_0　(b) a_4　(c) d_1　(d) d_2　(e) d_3　(f) d_4

例 7.3　用离散二进小波变换检测奇异性。

本例用表 7.1 所示滤波器组实现快速离散二进小波变换。为方便起见，滤波器系数均除以 $\sqrt{2}$，这相当于(7.3.27)式所表示的一组系数除以 $2^{j/2}$，这样(7.3.29)式和(7.3.30)式应修改成

$$|d_j(n)| \leqslant A(2^j)^{\alpha} \tag{7.3.31}$$

$$\log_2|d_j(n)| \leqslant \log_2 A + \alpha j \tag{7.3.32}$$

这实际上相当于用卷积型二进小波变换重新定义了 a_j 和 d_j。

图 7.8(a)为原始信号，它在 $n=100$ 和 200 处出现阶跃突变，在 300 处出现一尖脉冲。图 7.8(b)为经过 4 次平滑的信号 $a_4(n)$，它基本保持了信号的原貌。图 7.8(c)，(d)，(e)，(f)分别为 $j=1,2,3,4$ 时的离散二进小波变换 $d_1(n)$，$d_2(n)$，$d_3(n)$，$d_4(n)$。可以清楚地看到，在信号的奇异点处，离散二进小波变换的模确实出现了局部极大值。在发生阶跃突变的地方，模的局部极大值不随尺度变化，Lipschitz 指数为零。当奇异性为尖脉冲时，小波变换呈现为小波本身的波形。由于我们使用了图 7.6 所示具有双峰的样条小波，所以小波变换模的极大值有两个，分布在坐标为 300 附近，且随着尺度越来越精细，两个极大值的位置越来越靠近。由于此时 Lipschitz 指数为 -1，故模的极大值随 j 增加而减小。这再一次说明小波变换不仅可以检测信号奇异性发生的准确时间，而且可以进一步判断奇异的性质。

信号在 700，800 和 500 处的奇异性不是严格意义下的阶跃和冲击，这时可以认为信号 f 是具有严格奇异性的函数 f_0 和高斯函数的卷积积分，即

$$f(t) = f_0 * g_\sigma(t), \quad g_\sigma(t) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{t^2}{2\sigma^2})$$

设小波是某一平滑函数的1阶导数，则有

$$Wf(2^j,t) = 2^j \frac{d}{dt}(f_0 * g_\sigma * \bar{\theta}_j)(t)$$

假定

$$\bar{\theta}_j * g_\sigma(t) \approx \theta_{s'}(t), s' = \sqrt{2^{2j} + \sigma^2}$$

采用卷积型二进小波变换，则有

$$Wf(2^j,t) = 2^j \frac{d}{dt}(f_0 * \theta_{s'})(t) = \frac{2^j}{s'}Wf_0(s',t), \quad Wf_0(s',t) = f_0 * \psi_{s'}(t)$$

由于

$$|Wf_0(s',t)| \leqslant As'^{\alpha}$$

所以

$$|Wf(2^j,t)| \leqslant A2^j s'^{(\alpha-1)}, \quad s' = \sqrt{2^{2j} + \sigma^2}$$

由上式可以解释图7.6中信号在700,800和500处的小波变换模极大值随j变化的情况。用小波变换进行奇异性检测确实比常规的方法能获得更多的信息，但用Lipschitz指数判断奇异的性质还是比较困难的，需要工程技术人员把具体的工程问题和理论分析紧密地结合起来，反复试验，积累经验，这样才能取得较好的效果。

例7.4 一维a tuous算法。

a tuous算法和Mallat算法的主要区别在于：a tuous分解算法不需要抽取偶数样本，所以用于奇异检测时定位更准确；相应的，a tuous重构算法不需要插零。

若信号$x(n)$的持续时间为$0 \leqslant n \leqslant l_x - 1$，$FIR$滤波器$h(n)$的非零系数位于$0 \leqslant n \leqslant l_h - 1$，则输出$y(n)$的持续时间为$0 \leqslant n \leqslant l_h + l_x - 2$。也就是说，长度为$l_h$和$l_x$的两个有限长序列$h(n)$和$x(n)$卷积和的结果$y(n)$是一个长度为$l_h + l_x - 1$的有限长序列。调用信号处理工具箱函数conv实现两个一维数组的离散卷积，再将结果两端截断，得到和输入信号长度一致的输出，这样才能准确地确定奇异发生的时刻。本例采用表7.1的二进样条滤波器，程序如下：

```
% 装入文件名为“cuspamax”的信号 %
clear all;load cuspamax;s = cuspamax;length_of_s = length(s);
% 设置样条滤波器组系数 %
h = sqrt(2) * [0.125,0.375,0.375,0.125];g = sqrt(2) * [0.5, -0.5];
length_of_h = length(h);length_of_g = length(g);
% 设置分解级数，逼近和小波变换数组清零 %
J = 3;a(1:J+1,1:length_of_s) = 0;d(1:J+1,1:length_of_s) = 0;
% 第1级分解 %
sj = conv(s,h);wj = conv(s,g);
a(1,:) = sj(round((length_of_h + 1)/2):length(sj) - floor((length_of_h - 1)/2));
d(1,:) = wj(round((length_of_g + 1)/2):length(wj) - floor((length_of_g - 1)/2));
% 第2至J+1级分解 %
for j = 1:J
```

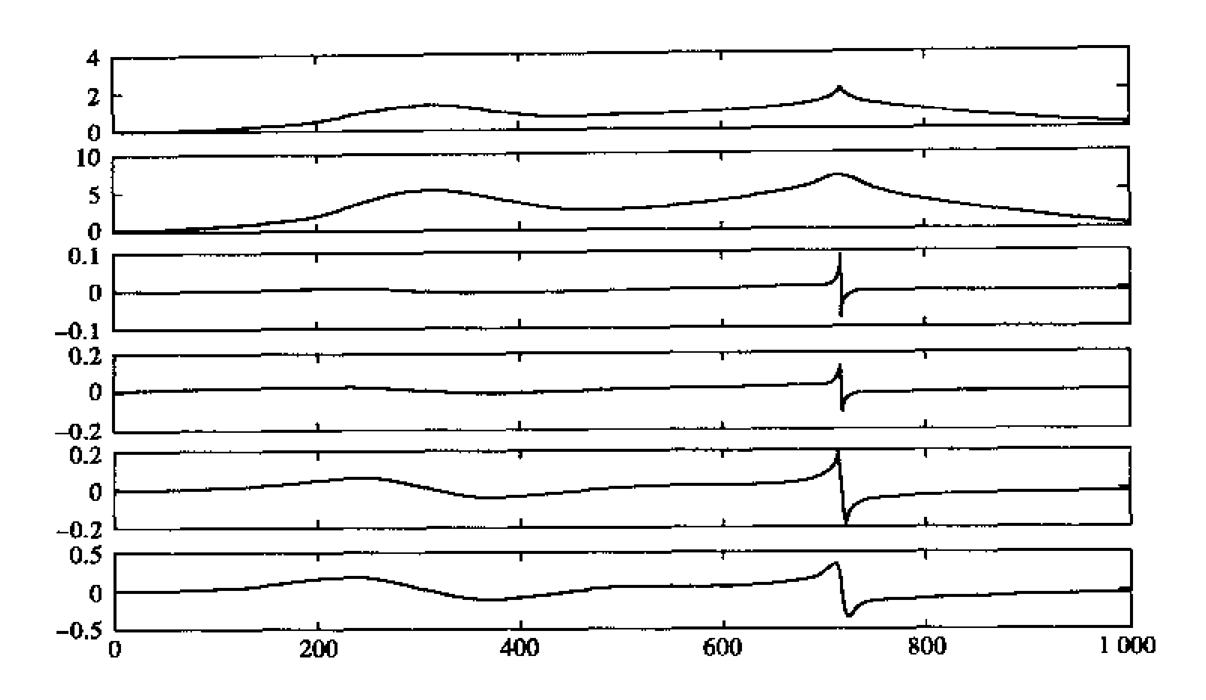


图7.9　一维二进小波变换

```
    length_of_hj = 2^j * (length_of_h - 1) + 1;
    length_of_gj = 2^j * (length_of_g - 1) + 1;
    hj(1:length_of_hj) = 0;
    gj(1:length_of_gj) = 0;
    for n = 1:length_of_h
        hj(2^j * (n - 1) + 1) = h(n);
    end
    for n = 1:length_of_g
        gj(2^j * (n - 1) + 1) = g(n);
    end
    sj = conv(a(j,:),hj);
    wj = conv(a(j,:),gj);
a(j+1,:) = sj(round((length_of_hj + 1)/2):length(sj) - floor((length_of_hj - 1)/2));
d(j+1,:) = wj(round((length_of_gj + 1)/2):length(wj) - floor((length_of_gj - 1)/2));
end
% 绘制原始信号、最后1级逼近信号和各级小波变换 %
subplot(J+3,1,1);plot(s)
subplot(J+3,1,2);plot(a(J+1,:))
for j = 1:J+1
subplot(J+3,1,j+2);plot(d(j,:))
end
```

程序运行结果示于图7.9，顶部为原始信号，往下依次为第4级平滑系数及第1到第4级二进小波变换系数。

7.4 图像多尺度边缘检测

物体的边缘表现为图像局部特性的不连续性,如灰度值的突变,颜色的突变。边缘常常意味着一个区域的终结和另一个区域的开始。图像边缘信息在图像分析和人的视觉中都是十分重要的,是图像识别中提取图像特征的一个重要属性。图像边缘有幅度和方向两个特性。通常,沿边缘走向的幅度变化平缓,垂直于边缘走向的幅度变化剧烈。此外,因物体大小不一,它们的边缘也有不同的尺度。Canny 边缘检测等价于在二维二进小波变换下检测模极大,边缘点的 Lipschitz 正则性取决于尺度细化过程中模极大的衰减速度。

7.4.1 二维二进小波变换

在二维情况下,Canny 边缘检测算法通过计算图像信号 $f(x,y)$ 的梯度矢量

$$\nabla f = \left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\right) \tag{7.4.1}$$

的模的局部极大值来寻找图像边缘的空间位置。梯度矢量的方向指出了图像灰度值变化最快的方向。

为了计算图像信号的两个偏导数,需要两个有方向性的二维小波,它们分别是二维平滑函数 $\theta(x,y)$ 的偏导数

$$\psi^x(x,y) = -\frac{\partial\theta(x,y)}{\partial x},\psi^y(x,y) = -\frac{\partial\theta(x,y)}{\partial y} \tag{7.4.2}$$

$\theta(x,y)$ 在 $x-y$ 平面的积分为 1,且很快地收敛到零。

令

$$\psi_j^x(x,y) = 2^{-j}\psi^x(2^{-j}x,2^{-j}y),\psi_j^y(x,y) = 2^{-j}\psi^y(2^{-j}x,2^{-j}y) \tag{7.4.3}$$

并定义小波变换的两个分量

$$\begin{aligned} W^x f(2^j,x,y) &= (f(u,v),\psi_j^x(u-x,v-y)) = f*\overline{\psi}_j^x(x,y) \\ W^y f(2^j,x,y) &= (f(u,v),\psi_j^y(u-x,v-y)) = f*\overline{\psi}_j^y(x,y) \end{aligned} \tag{7.4.4}$$

其中,

$$\overline{\psi}_j^x(x,y) = \psi_j^x(-x,-y),\overline{\psi}_j^y(x,y) = \psi_j^y(-x,-y) \tag{7.4.5}$$

任意 $f\in L^2(R^2)$ 的二进小波变换定义为如下函数族

$$Wf(2^j,x,y) = \{W^x f(2^j,x,y),W^y f(2^j,x,y)\}_{j\in Z} \tag{7.4.6}$$

为确保二进小波变换的完备性和稳定性,必须满足如下充分必要条件:存在两个正常数 A 和 B,对 $\forall(\omega_x,\omega_y)\in R^2-\{(0,0)\}$ 使

$$A \leqslant \sum_{j=-\infty}^{\infty}|\hat{\psi}^x(2^j\omega_x,2^j\omega_y)|^2 + |\hat{\psi}^y(2^j\omega_x,2^j\omega_y)|^2 \leqslant B \tag{7.4.7}$$

其中,$\hat{\psi}^x$ 和 ψ^y 分别表示 ψ^x 和 ψ^y 的二维傅立叶变换。满足上式的 $\{\psi^x,\psi^y\}$ 称为二进小波,对二进小波,存在重构小波 $\{\tilde{\psi}^x,\tilde{\psi}^y\}$,它们的傅立叶变换满足

$$\sum_{j=-\infty}^{\infty}2^{-2j}[\hat{\tilde{\psi}}^x(2^j\omega_x,2^j\omega_y)\hat{\psi}^{x*}(2^j\omega_x,2^j\omega_y) + \hat{\tilde{\psi}}^y(2^j\omega_x,2^j\omega_y)\hat{\psi}^{y*}(2^j\omega_x,2^j\omega_y)] = 1 \tag{7.4.8}$$

因而

$$f(x,y) = \sum_{j=-\infty}^{\infty} 2^{-2j}[W^x f(2^j,x,y) * \tilde{\psi}_j^x(x,y) + W^y f(2^j,x,y) * \tilde{\psi}_j^y(x,y)] \quad (7.4.9)$$

7.4.2　多尺度边缘检测

与一维情况类似，由于$\{\psi^x,\psi^y\}$是平滑函数$\theta(x,y)$的1阶偏导数，所以二维二进小波变换的两个分量等价于信号$f(x,y)$被平滑后的梯度矢量的两个分量，即

$$\begin{pmatrix} W^x f(2^j,x,y) \\ W^y f(2^j,x,y) \end{pmatrix} = 2^j \begin{pmatrix} \frac{\partial}{\partial x}(f * \bar{\theta}_j)(x,y) \\ \frac{\partial}{\partial y}(f * \bar{\theta}_j)(x,y) \end{pmatrix} = 2^j \nabla (f * \bar{\theta}_j)(x,y) \quad (7.4.10)$$

梯度矢量$\nabla(f * \bar{\theta}_j)(x,y)$的模正比于

$$Mf(2^j,x,y) = \sqrt{|W^x f(2^j,x,y)|^2 + |W^y f(2^j,x,y)|^2} \quad (7.4.11)$$

而梯度矢量与水平方向的夹角为

$$Af(2^j,x,y) = \begin{cases} \alpha(x,y), & 如\ W^x f(2^j,x,y) \geqslant 0 \\ \pi - \alpha(x,y), & 如\ W^x f(2^j,x,y) < 0 \end{cases} \quad (7.4.12)$$

其中，

$$\alpha(x,y) = \tan^{-1}\left(\frac{W^y f(2^j,x,y)}{W^x f(2^j,x,y)}\right) \quad (7.4.13)$$

用二进小波变换实现多尺度边缘检测就是寻找$Mf(2^j,x,y)$的局部极大值，$Af(2^j,x,y)$指明了边缘的方向。除确定边缘的位置和方向外，还可以用小波变换的衰减速度判断边缘的奇异性。对Lipschitz指数$0\leqslant\alpha\leqslant1$，如存在常数$A>0$，对所有的$(x,y)\in R^2$，使得

$$|f(x,y) - f(x_0,y_0)| \leqslant A(|x-x_0|^2 + |y-y_0|^2)^{\alpha/2} \quad (7.4.14)$$

则称函数f在(x_0,y_0)点Lipschitz α。如对区域$(x_0,y_0)\in\Omega$内的所有点，都存在$A>0$，使上式成立，则称函数f在Ω内一致Lipschitz α。与一维情况的定理7.2类似，可以证明：当且仅当存在$A>0$，对所有尺度2^j及区域Ω内的所有点，使得

$$|Mf(2^j,x,y)| \leqslant A2^{j(\alpha+1)} \quad (7.4.15)$$

则f在Ω内一致Lipschitz α。

7.4.3　快速多尺度边缘检测算法

作边缘检测的二维二进小波可以设计为一维二进小波的可分积，具体地说，它们的傅立叶变换表示为

$$\hat{\psi}^x(\omega_x,\omega_y) = G(w_x/2)\hat{\phi}(\omega_x/2)\hat{\phi}(\omega_y/2) \quad (7.4.16)$$

$$\hat{\psi}^y(\omega_x,\omega_y) = G(w_y/2)\hat{\phi}(\omega_x/2)\hat{\phi}(\omega_y/2) \quad (7.4.17)$$

其中，$\hat{\phi}(\omega)$是一个低通滤波器，而

$$G(\omega) = -i\sqrt{2}e^{-i\omega/2}\sin(\omega/2) \quad (7.4.18)$$

是一个高通数字滤波器。

为了能用滤波器组快速实现二维离散二进小波变换，与一维情况一样，假定尺度函数满足如下二尺度方程

$$\hat{\phi}(\omega) = \prod_{p=1}^{+\infty} \frac{H(2^{-p}\omega)}{\sqrt{2}} = \frac{1}{2}H\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \tag{7.4.19}$$

若选择尺度函数为 m 次样条,即

$$\hat{\phi}(\omega) = e^{-\frac{i\varepsilon\omega}{2}}\left(\frac{\sin(\omega/2)}{(\omega/2)}\right)^{m+1}, \varepsilon = \begin{cases} 0, & \text{当 } m \text{ 为奇数} \\ 1, & \text{当 } m \text{ 为偶数} \end{cases} \tag{7.4.20}$$

则可得

$$H(\omega) = \sqrt{2}e^{-i\varepsilon\omega/2}[\cos(\omega/2)]^{m+1} \tag{7.4.21}$$

与一维情况一样,对二进小波变换在所有尺度时都均匀采样,假定采样间隔等于 1,则离散小波系数为

$$d_j^x(n,m) = W^x(2^j,n,m), d_j^y(n,m) = W^y(2^j,n,m) \tag{7.4.22}$$

同样定义原始图像信号

$$a_0(n,m) = \langle f(x,y), \phi(x-n)\phi(y-m) \rangle \tag{7.4.23}$$

和 $j \geq 0$ 时的平滑图像信号

$$a_j(n,m) = \langle f(x,y), \phi_j(x-n)\phi_j(y-m) \rangle \tag{7.4.24}$$

那么,二维离散二进小波变换的 à trous 算法表示为如下离散卷积形式

$$a_{j+1}(n,m) = a_j * \bar{h}_j\bar{h}_j(n,m) \tag{7.4.25}$$

$$d_{j+1}^x(n,m) = a_j * \bar{g}_j\delta(n,m) \tag{7.4.26}$$

$$d_{j+1}^y(n,m) = a_j * \delta\bar{g}_j(n,m) \tag{7.4.27}$$

其中,

$$\bar{h}_j\bar{h}_j(n,m) = \bar{h}_j(n)\bar{h}_j(m) \tag{7.4.28}$$

$$\bar{g}_j\delta(n,m) = \bar{g}_j(n)\delta(m) \tag{7.4.29}$$

$$\delta\bar{g}_j(n,m) = \delta(n)\bar{g}_j(m) \tag{7.4.30}$$

也就是说,a_{j+1} 是 a_j 沿横向和纵向低通滤波的结果,而 d_{j+1}^x 是 a_j 沿横向高通滤波的结果,d_{j+1}^y 是 a_j 沿纵向高通滤波的结果。

例 7.5　二维 a tuous 算法。

设灰度图像 $x(n,m)$ 为 N_x 行 $\times M_x$ 列矩阵,二维数字滤波器 $h(n,m)$ 为 N_h 行 $\times M_h$ 列矩阵,与一维情况类似,输出 $y(n,m)$ 为 (N_x+N_h-1) 行 $\times(M_x+M_h-1)$ 列矩阵。现在用图像处理工具箱函数 conv2 实现灰度图像 $x(n,m)$ 和二维数字滤波器 $h(n,m)$ 的离散卷积,调用 conv2 时对 shape 参数赋字符串 'same',得到的输出图像 $y(n,m)$ 与输入图像 $x(n,m)$ 具有相同的大小。

必须注意,灰度图像经过数字滤波等数值计算后,矩阵元素可能出现负值。例如经过高通滤波,一般都会出现负值,因为高通滤波体现了灰度的变化,如灰度由暗变亮时出现正值,则由亮变暗时就会出现负值,反之亦然。在进行图像边缘检测时,主要关心的是小波变换的模,所以这个例子也主要是计算小波变换的模。程序如下:

```
% 装入文件名为"wbarb"的索引图像,并将其转换为灰度图像 %
clear all;load wbarb;I = ind2gray(X,map);
I = imadjust(I,stretchlim(I),[0 1]); [N,M] = size(I);subplot(2,2,1);imshow(I);
% 设置样条滤波器系数 %
```

```
h = [0.125,0.375,0.375,0.125];g = [0.5, -0.5];delta = [1,0,0];
% 设置分解级数,逼近和 x,y 方向的二进小波系数及梯度绝对值数组清零 %
J = 2;
a(1:N,1:M,1:J+1) = 0;
dx(1:N,1:M,1:J+1) = 0;
dy(1:N,1:M,1:J+1) = 0;
d(1:N,1:M,1:J+1) = 0;

% 第 1 级分解,显示第 1 级分解的边缘 %
a(:,:,1) = conv2(h,h,I,'same');
dx(:,:,1) = conv2(delta,g,I,'same');
dy(:,:,1) = conv2(g,delta,I,'same');
x = dx(:,:,1);
y = dy(:,:,1);
d(:,:,1) = sqrt(x.^2 + y.^2);
I = imadjust(d(:,:,1),stretchlim(d(:,:,1)),[0 1]);subplot(2,2,2);imshow(I);

% 第 2 至 J+1 级分解 %
lh = length(h);
lg = length(g);
for j = 1:J
    lhj = 2^j * (lh - 1) + 1;
    lgj = 2^j * (lg - 1) + 1;
    hj(1:lhj) = 0;
    gj(1:lgj) = 0;
    for n = 1:lh
        hj(2^j * (n - 1) + 1) = h(n);
    end
    for n = 1:lg
        gj(2^j * (n - 1) + 1) = g(n);
    end
    a(:,:,j+1) = conv2(hj,hj,a(:,:,j),'same');
    dx(:,:,j+1) = conv2(delta,gj,a(:,:,j),'same');
    dy(:,:,j+1) = conv2(gj,delta,a(:,:,j),'same');
    x = dx(:,:,j+1);
    y = dy(:,:,j+1);
    d(:,:,j+1) = sqrt(x.^2 + y.^2);
    I = imadjust(d(:,:,j+1),stretchlim(d(:,:,j+1)),[0 1]);subplot(2,2,j+2);imshow(I);
```

```
end
```

程序运行结果如图7.10所示，左上角为原始图像，从右上角到左下角、右下角依此为第1,2,3级小波变换的模，第1级小波变换模显示出图像的边缘和纹理（如头巾），第2和第3级小波变换模则主要显示出图像的边缘，平滑掉了图像细致的纹理结构。






图7.10　图像二进小波变换

第 8 章

逼近、去噪及数据压缩

按照逼近论的思想，逼近是选择最能刻画信号固有性质的正交基，在变换域用尽可能少的系数去描述信号。信号逼近最直接的应用领域是数据压缩和去噪。选择最能刻画信号固有性质的正交基，实际上就是选择了一种描述方式，使信号和噪声在最大限度上区分开来，从而实现噪声的滤除。一般来说，信号在时间域内有很强的相关性，图像信号相邻像素之间，运动图像相邻帧之间都有很强的相关性，在适当选择的正交基下，变换域系数之间的相关性大大减弱，这是变换编码实现数据压缩的根本原因。

8.1 逼　近

数值逼近试图选择适当的正交基，对信号进行分解，使得只有少量分解系数有较大幅值，从而可以用它们去有效地逼近信号。这时信号能量聚集在少数基矢量上，逼近误差体现为信号分布在其余基矢量上的能量。信号的逼近实际上就是它在这些基矢量张成的子空间上的投影。如先验地选定正交基的前 M 个去逼近信号，称为线性逼近，小波多分辨率逼近就是一种线性逼近。线性逼近的误差随 M 增加而衰减的速度取决于信号的正则性，比较适宜于较平滑信号的逼近。非线性逼近则根据分解系数的幅值自适应地选择基矢量，小波阈值逼近就是一种非线性逼近，算法也很简单，很适合分段正则信号的逼近。在非线性逼近的基础上，提出了选取最优小波包基的准则和相应的算法。

8.1.1 线性逼近

一般来说，线性逼近比较简单，它选择正交基的前 M 个去逼近信号。设 $B=\{g_m\}_{m\in N}$是 Hilbert 空间的正交规范基，m 为变换域指标，g_m 是连续时间函数或离散时间序列，任意信号 $f\in H$可分解为

$$f=\sum_{m=0}^{+\infty}\langle f,g_m\rangle g_m \tag{8.1.1}$$

如事先选择正交基的 M 个系数，例如选择正交基的前 M 个系数去逼近信号，即

$$f_M = \sum_{m=0}^{M-1} \langle f, g_m \rangle g_m \tag{8.1.2}$$

则逼近的均方误差可用剩余系数的范数表示

$$\varepsilon_l(M) = \|f - f_M\|^2 = \sum_{m=M}^{+\infty} |\langle f, g_m \rangle|^2 \tag{8.1.3}$$

由于$\|f\|^2 = \sum_{m=0}^{+\infty} |\langle f, g_m \rangle|^2 < +\infty$，所以随 M 的增加，逼近误差逐渐衰减到 0，即

$$\lim_{M \to +\infty} \varepsilon_l(M) = 0 \tag{8.1.4}$$

衰减速度取决于信号的正则性。

对傅立叶分析而言，线性逼近意味着用较低频率的正弦波去逼近信号。傅立叶系数的衰减速度取决于信号的正则性，信号越正则傅立叶系数衰减越快，对同样的 M 逼近误差越小；或者说，在给定逼近误差的条件下，对比较正则的信号只需要用较少的低频系数去逼近。

对小波分析而言，线性逼近实际上就是多分辨率逼近。对任意$f \in L^2(R)$，有

$$f = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \tag{8.1.5}$$

信号在分辨率为2^{-L}时的逼近实际上就是信号在 V_L 上的投影，由于

$$V_L = W_{L+1} \oplus V_{L+1} = W_{L+1} \oplus W_{L+2} \oplus V_{L+2} = \cdots = \bigoplus_{j=L+1}^{+\infty} W_j$$

所以逼近信号可写成

$$f_{A_L} = \sum_{j=L+1}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} = \sum_{n=-\infty}^{+\infty} \langle f, \phi_{L,n} \rangle \phi_{L,n} \tag{8.1.6}$$

也就是说，逼近信号是所有较粗尺度上细节信号的总和。由上列两式得到逼近误差

$$\varepsilon_l(L) = \|f - f_{A_L}\|^2 = \sum_{j=-\infty}^{L} \sum_{n=-\infty}^{+\infty} |\langle f, \psi_{j,n} \rangle|^2$$

它是所有细尺度下小波系数的能量。显然，小波系数的衰减速度也取决于信号的正则性。对正则性较好的信号，可以用较大的尺度去逼近，对正则性较差的信号，较精细尺度下的小波系数才能刻画其小范围内的快速变化，即要用较精细的尺度去逼近。这与傅立叶逼近实质上是一致的，较高的频率意味着更精细的尺度。

8.1.2 *K-L* 基

对具有某些共同性质的信号的全体，可以用随机过程信号模型来描述它们。例如语音信号可以看成高斯随机过程。假定信号可以用零均值随机过程模型 $F(n)$ 来描述，信号的一次实现有 N 个样本，则随机矢量在正交规范基$\{g_m\}_{0 \le m \le N-1}$下分解为

$$F = \sum_{m=0}^{N-1} \langle F, g_m \rangle g_m = \sum_{m=0}^{N-1} F_B(m) g_m \tag{8.1.7}$$

其中，

$$F_B(m) = \langle F, g_m \rangle = \sum_{n=0}^{N-1} F(n) g_m^*(n) \tag{8.1.8}$$

是信号在变换域内的分解系数。

如只选择正交基的前 M 个系数去逼近信号，即

$$F_M = \sum_{m=0}^{M-1} \langle F, g_m \rangle g_m \tag{8.1.9}$$

则逼近的均方误差可用剩余系数的范数表示

$$\varepsilon_l(M) = E\{\|F - F_M\|^2\} = \sum_{m=M}^{+\infty} E\{|\langle F, g_m \rangle|^2\} \tag{8.1.10}$$

其中,

$$E\{|\langle F, g_m \rangle|^2\} = E\{\langle F, g_m \rangle \langle F, g_m \rangle^*\} = E\left\{\sum_{n=0}^{N-1} F(n) g_m^*(n) \sum_{k=0}^{N-1} F^*(k) g_m(k)\right\}$$
$$= \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} E\{F(n)F^*(k)\} g_m(k) g_m^*(n) = \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} R(n,k) g_m(k) g_m^*(n)$$

其中,$R(n,k) = E\{F(n)F^*(k)\}$是随机过程的自相关矩阵。满足

$$\sum_{k=0}^{N-1} R(n,k) g_m(k) = \sigma_m^2 g_m(n) \tag{8.1.11}$$

的正交基称为 *K-L* 基。也就是说 *K-L* 基的每个矢量 g_m 都是自相关矩阵的特征矢量,σ_m^2 为对应的特征值。对 *K-L* 基而言,$E\{|\langle F, g_m \rangle|^2\} = \sigma_m^2$,故线性逼近的均方误差为

$$\varepsilon_l(M) = E\{\|f - f_M\|^2\} = \sum_{m=M}^{N-1} \sigma_m^2 \tag{8.1.12}$$

不难理解,如将 g_m 按特征值 σ_m^2 递减的顺序排列,逼近误差最小。

定理 8.1 当且仅当$\{g_m\}_{0 \le m \le N-1}$是 *K-L* 基,且特征矢量按特征值递减顺序排列时,线性逼近的均方误差最小。

K-L 基还有一个十分重要的性质。显然,内积系数 $F_B(m)$ 是随机变量,对零均值随机信号 $F(n)$,$F_B(m)$的均值为 0。在 *K-L* 基下,内积系数的互相关是

$$E\{F_B(m)F_B^*(p)\} = E\left\{\sum_{n=0}^{N-1} F(n) g_m^*(n) \sum_{k=0}^{N-1} F^*(k) g_p(k)\right\}$$
$$= \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} E\{F(n)F^*(k)\} g_p(k) g_m^*(n)$$
$$= \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} R(n,k) g_p(k) g_m^*(n) = \sum_{n=0}^{N-1} \sigma_p^2 g_p(n) g_m^*(n)$$

于是有

$$E\{F_B(m)F_B^*(p)\} = \sigma_p^2 \delta(p - n) \tag{8.1.13}$$

也就是说,在 *K-L* 基下,信号在变换域内的分解系数是不相关的。这是一个十分重要的结论。

例 8.1 若 $F(n)$是一个广义平稳零均值随机过程,则其自相关是时间差的函数,即

$$R(n,k) = E\{F(n)F^*(k)\} = R(n - k)$$

如 $R(n)$是以 N 为周期的,则称为循环平稳过程,(8.1.11)式是一个循环卷积表达式。而循环卷积的特征矢量就是离散傅立叶矢量

$$\left\{g_k(n) = \frac{1}{\sqrt{N}} \exp\left(\frac{i2\pi kn}{N}\right)\right\}_{0 \le k \le N-1}$$

因而循环平稳过程的 *K-L* 基正是离散傅立叶基,而特征值

$$\sigma_k^2 = \hat{R}(k) = \sum_{n=0}^{N-1} R(n) \exp\left(\frac{-i2\pi kn}{N}\right)$$

正是随机过程的功率谱。

8.1.3 小波阈值逼近

对正则信号,傅立叶逼近和小波逼近具有相同的效率。但对有若干孤立奇异点的分段正则信号,傅立叶变换需要用大量的高频正弦波去逼近信号的奇异点,逼近误差衰减的速度很慢,傅立叶分析的效率很低。而这时只有少量小波系数幅度较大,故可选择这些幅度较大的小波系数去逼近信号,这是一种非线性逼近。

线性逼近是针对一类信号事先选定正交基,从中再选择 M 个基矢量,信号在这 M 个基矢量张成的子空间上的正交投影作为逼近。线性逼近与信号无关,逼近算子是一个线性算子。非线性逼近则是针对具体信号选取使内积系数较大的 M 个矢量,记 I_M 为这 M 个矢量的指标集,则信号 f 的非线性逼近为

$$f_M = \sum_{m \in I_M} \langle f, g_m \rangle g_m \tag{8.1.14}$$

逼近的均方误差为

$$\varepsilon(M) = \|f - f_M\|^2 = \sum_{m \notin I_M} |\langle f, g_m \rangle|^2 \tag{8.1.15}$$

它实际上是残存在幅度较小的内积系数上的能量。这时指标集因信号不同而不同,不能事先选定,所以是非线性的。非线性逼近依赖于具体信号的性质,具有自适应性。

选择幅度较大的系数实际上是用一个阈值 T 去决定所有内积系数的取舍,表示为

$$\theta_T(x) = \begin{cases} x, & 若\ |x| \geq T \\ 0, & 若\ |x| < T \end{cases} \tag{8.1.16}$$

于是信号 f 的阈值逼近表示为

$$f_M = \sum_{m=0}^{+\infty} \theta_T(\langle f, g_m \rangle) g_m \tag{8.1.17}$$

显然,阈值逼近的矢量指标集因信号而异,所以它是一种非线性逼近。

对正交小波基而言,信号可分解成

$$f = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \tag{8.1.18}$$

阈值逼近

$$f_M = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \theta_T(\langle f, \psi_{j,n} \rangle) \psi_{j,n} \tag{8.1.19}$$

它实际上是保留那些幅度大于或等于阈值的小波系数,而将幅度小于阈值的小波系数置为0。因为小波系数的幅度与信号的局部奇异性有关,所以对只有少数孤立奇异点的分段正则信号而言,幅度大的小波系数很少,阈值逼近比线性逼近更有效率。选择大幅度的小波系数意味着在信号奇异点附近,选用较精细的尺度,而在光滑部分,则用较粗糙的尺度,这等价于在时间-尺度平面上构造了一个与信号性质自适应的网格,而线性逼近则是在整个信号支集保持相同的分辨率。此外,阈值算法非常简单,所以阈值逼近在小波分析中获得了广泛应用。

8.1.4 最优小波包逼近

小波包分析为逼近提供了数量庞大的可供选择的正交基,称之为一个字典 D。D 其实是

若干可供选择的正交基的并集：

$$D = \bigcup_{\lambda \in \Lambda} B^{\lambda} \tag{8.1.20}$$

对 N 维信号空间，每一组正交基都由 N 个矢量构成

$$B^{\lambda} = \{g_m^{\lambda}\}_{1 \leqslant m \leqslant N} \tag{8.1.21}$$

在基 B^{λ} 中的非线性逼近为

$$f_M^{\lambda} = \sum_{m \in I_M^{\lambda}} \langle f, g_m^{\lambda} \rangle g_m^{\lambda} \tag{8.1.22}$$

逼近的均方误差为

$$\varepsilon^{\lambda}(M) = \|f - f_M^{\lambda}\|^2 = \sum_{m \notin I_M^{\lambda}} |\langle f, g_m^{\lambda} \rangle|^2 \tag{8.1.23}$$

选择适当的指标集 I_M^{λ} 使误差最小的逼近称为在基 B^{λ} 中的最佳逼近。上面已经讨论了在一个确定的基 B^{λ} 中的逼近。而对于小波包逼近来说，还有一个在基字典 D 中查找最优基的问题。为此，先要对如何比较基字典中两个基的优劣下一个确切的定义。

定义 8.1　称 $B^{\alpha} = \{g_m^{\alpha}\}_{1 \leqslant m \leqslant N}$ 是比 $B^{\beta} = \{g_m^{\beta}\}_{1 \leqslant m \leqslant N}$ 更好的基，当且仅当对所有的 $M \geqslant 1$

$$\varepsilon^{\alpha}(M) \leqslant \varepsilon^{\beta}(M) \tag{8.1.24}$$

上式必须对所有可能的 M 值成立，才能说 B^{α} 是比 B^{β} 更好的基。如存在 M_0 和 M_1 使得

$$\varepsilon^{\alpha}(M_0) < \varepsilon^{\beta}(M_0) \quad 且 \quad \varepsilon^{\alpha}(M_1) > \varepsilon^{\beta}(M_1) \tag{8.1.25}$$

则不能判定 B^{α} 和 B^{β} 两者之中哪一个是较好的基。

倘若要在一个数量庞大的字典中搜寻最优基，那将是一件十分繁杂的工作。为此，先引入花费。在基 B^{λ} 中的花费定义为

$$C(f, B^{\lambda}) = \sum_{m=1}^{N} \Phi\left(\frac{|\langle f, g_m^{\lambda} \rangle|^2}{\|f\|^2}\right) \tag{8.1.26}$$

其中，$\Phi(x)$ 是一个 Schur 凹函数。花费是对归一化的内积系数的范数来定义的，显然应要求 $\Phi(0) = 0$，也就是说，如某一内积系数等于 0，那么它就没有花费；而且花费是可加的。下述定理进一步将花费和基的优劣联系起来。

定理 8.2　当且仅当对所有的凹函数 $\Phi(x)$

$$\sum_{m=1}^{N} \Phi\left(\frac{|\langle f, g_m^{\alpha} \rangle|^2}{\|f\|^2}\right) \leqslant \sum_{m=1}^{N} \Phi\left(\frac{|\langle f, g_m^{\beta} \rangle|^2}{\|f\|^2}\right) \tag{8.1.27}$$

则 B^{α} 是比 B^{β} 更好的基。

这个定理意味着，可以用花费的多少来判断基的优劣，用花费少的基作非线性逼近比花费多的基所带来的误差小。在变换域中，若大幅度内积系数只集中在少量基矢量 g_m^{λ} 上，则花费少，非线性逼近误差小，刻画信号本质特性的效率高。让我们考虑两个最极端的情况。如信号只需要一个基矢量来表征，即 $f = \eta g_k^{\lambda}$，其中 η 为一常数，则 $C(f, B^{\lambda}) = \Phi(1)$，这时花费最少。相反，如信号需要所有基矢量来表征，且信号能量均匀地分布在所有基矢量上，即

$$|\langle f, g_m^{\lambda} \rangle|^2 = \frac{\|f\|^2}{N}, \qquad 1 \leqslant m \leqslant N$$

则花费为 $C(f, B^{\lambda}) = N\Phi(N^{-1})$。可以证明[3]：任意一组基的花费都介于上述两种极端情况之间，即

$$\Phi(1) \leqslant C(f, B^{\lambda}) \leqslant N\Phi(N^{-1}) \tag{8.1.28}$$

至于具体的花费函数形式，已在6.3节提及，不再赘述。

在实际应用中，往往只用一个花费函数来比较基的优劣。定理8.2证明：如 B^{α} 是比 B^{β} 更好的基，则对所有的花费函数，都有

$$C(f,B^{\alpha}) \leqslant C(f,B^{\beta}) \tag{8.1.29}$$

如上式仅对某一个花费函数成立，则不能断定 B^{α} 肯定是比 B^{β} 更好的基。Coifman 和 Wiicker-hauser 通过最小化花费寻找字典 D 中的最好基 B^{α}

$$C(f,B^{\alpha}) = \min_{\lambda \in \Lambda} C(f,B^{\lambda}) \tag{8.1.30}$$

并证明 D 中不存在比 B^{α} 更好的基。然而，常常会有其他的基如 B^{γ}，使得

$$\varepsilon^{\alpha}(M_0) < \varepsilon^{\gamma}(M_0) \text{ 及 } \varepsilon^{\alpha}(M_1) > \varepsilon^{\gamma}(M_1) \tag{8.1.31}$$

根据好基的定义，不能说 B^{α} 更好，但也不能说 B^{γ} 更好，只能说它们在(8.1.31)式的意义下是等价的。显然，用一个花费函数最小化花费的方法寻找最优基在实际应用中比较可行，但这样得到的最优基取决于花费函数的选择。关于最优小波包基的搜索算法，6.3和6.4节已作了详细说明，不再赘述。

8.2 去 噪

8.2.1 对角估计

设信号 f 被加性噪声 W 污染，则观测信号表示为

$$X = f + W \tag{8.2.1}$$

从统计学的观点看，去噪是用实际上可观测到的数据 X 对信号 f 的估计。下面讨论变换域内的滤波问题。在正交规范基 $B = \{g_m\}_{0 \leqslant m \leqslant N-1}$ 下观测信号 X 被分解为

$$X_B(m) = f_B(m) + W_B(m) \tag{8.2.2}$$

显然，

$$X_B(m) = \langle X,g_m \rangle, f_B(m) = \langle f,g_m \rangle, W_B(m) = \langle W,g_m \rangle \tag{8.2.3}$$

假设 W 是方差为 σ^2 的零均值白噪声，即

$$E[W(n)W^*(k)] = \sigma^2 \delta(n-k)$$

则 $E[W_B(m)] = 0$，且

$$\begin{aligned} E[W_B(m)W_B^*(p)] &= \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} g_m^*(n) g_p(k) E[W(n)W^*(k)] \\ &= \sigma^2 \sum_{n=0}^{N-1} g_m^*(n) g_p(n) = \sigma^2 \delta(p-m) \end{aligned}$$

也就是说，方差为 σ^2 的零均值白噪声在变换域的内积系数仍然是方差为 σ^2 的零均值白噪声，所以白噪声的能量在变换域是均匀分布的。

由于在变换域内信号能量集中在少量大幅值内积系数上，而噪声能量在变换域却是均匀分布的，所以可从一个特定的 $X_B(m)$ 来独立地估计 $f_B(m)$，从而得到信号的估计

$$\tilde{F} = DX = \sum_{m=0}^{N-1} d_m \left(X_B(m) \right) g_m \tag{8.2.4}$$

其中，D 称为对角估计算子。因为噪声均值为零，所以当观测信号为零时，估计信号为零显然是合理的，故 $D0=0$，因而 $d_m(0)=0$，于是对角估计函数可以表示为

$$d_m\left(X_B(m)\right)=\alpha(m)X_B(m),\quad 0\leqslant m\leqslant N-1 \tag{8.2.5}$$

估计的均方误差称为风险，它可以用内积系数表示为

$$r(D,f)=E\{\|f-\tilde{F}\|^2\}=\sum_{m=0}^{N-1}E\{|f_B(m)-\alpha(m)X_B(m)|^2\} \tag{8.2.6}$$

线性估计假定信号能量集中在前 M 个系数上，于是当 $0\leqslant m\leqslant M-1$ 时，选择 $\alpha(m)=1$，而当 $M\leqslant m\leqslant N-1$ 时，$\alpha(m)=0$，由(8.2.6)得线性估计的风险为

$$r(D,f)=M\sigma^2+\sum_{m=M}^{N-1}|f_B(m)|^2=M\sigma^2+\varepsilon_l(M) \tag{8.2.7}$$

上式第 1 项为噪声带来的估计误差，第 2 项为线性逼近误差。M 越大，线性逼近误差越小，但噪声带来的估计误差越大；反之，M 越小，噪声带来的估计误差越小，但线性逼近误差越大。所以应调整 M 使这两项具有相同数量级。显然，线性估计不是最优的，尤其对非正则信号更是如此。

假定信号和噪声不相关，可推得

$$E\{|f_B(m)-\alpha(m)X_B(m)|^2\}=|f_B(m)|^2\left(1-\alpha(m)\right)^2+\sigma^2\alpha_m^2(m)$$

为了得到风险最小的对角估计算子，将上式对 $\alpha(m)$ 求导并令其为零得

$$\alpha(m)=\frac{|f_B(m)|^2}{|f_B(m)|^2+\sigma^2} \tag{8.2.8}$$

这个结果是合理的：当信噪比 $|f_B(m)|^2/\sigma^2$ 很大时，$\alpha(m)\approx1$，$f_B(m)$ 的估计为 $X_B(m)$；当信噪比很小时，$\alpha(m)\approx0$，估计 $f_B(m)$ 为 0；在其他情况下，则根据信噪比的大小对 $X_B(m)$ 按比例进行衰减作为 $f_B(m)$ 的估计。将(8.2.8)式代入(8.2.6)式得最小风险

$$r_{\inf}(f)=\sum_{m=0}^{N-1}\frac{|f_B(m)|^2\sigma^2}{|f_B(m)|^2+\sigma^2} \tag{8.2.9}$$

在有噪声污染的情况下，实际上只能得到观测信号，$f_B(m)$ 是未知的，所以不能计算 $\alpha(m)$，$r_{\inf}(f)$ 只是一个实际上不能达到的理论上的最小风险。

Oracle 投影法是一种非线性估计，它用噪声电平作为阈值，对信号的内积系数进行判别，以决定观测信号内积系数的取舍，即选择

$$\alpha(m)=\begin{cases}1, & 若\ |f_B(m)|\geq\sigma\\ 0, & 若\ |f_B(m)|<\sigma\end{cases} \tag{8.2.10}$$

这实际上是对上述最优估计的一种简化，当信噪比大于 1 时，用 $X_B(m)$ 作为 $f_B(m)$ 的估计，而当信噪比小于 1 时，估计 $f_B(m)$ 为 0。由(8.2.6)式可得 Oracle 投影的估计风险为

$$r_p(f)=\sum_{m=0}^{N-1}\min[\,|f_B(m)|^2,\sigma^2\,] \tag{8.2.11}$$

设 M 个系数是信噪比大于 1，则上式可写成

$$r_p(f)=M\sigma^2+\sum_{|f_B(m)|<\sigma}|f_B(m)|^2=M\sigma^2+\varepsilon_n(M) \tag{8.2.12}$$

其中第 1 项为噪声带来的估计误差，第 2 项为非线性逼近误差。为使风险减小，应找到信号的稀疏表示，使信号能量聚集在少数系数上，这样 M 很小且 $\varepsilon_n(M)$ 也比较小。在 M 相同的情况

下,非线性逼近误差总是小于线性逼近误差,所以 Oracle 投影估计的误差要小于线性估计的误差。

由于对任意 x,y 有

$$\min(x,y) \geqslant \frac{xy}{x+y} \geqslant \frac{1}{2}\min(x,y)$$

所以

$$r_p(f) \geqslant r_{\inf}(f) \geqslant \frac{1}{2}r_p(f) \tag{8.2.13}$$

也就是说,Oracle 投影估计风险已经非常接近理论极限,它是风险最小的非线性估计。

8.2.2 阈值估计

上述非线性估计实际上是一种阈值估计,或者说是一种幅值域的滤波方法。它根据信号系数$f_B(m)$的大小决定观测数据的取舍。但由于$f_B(m)$是未知的,所以还是不可计算的。对它最简单的改进是:不再用信号内积系数进行阈值判别,而直接用观测信号内积系数进行阈值判别,这是实际可行的一种简单算法。而且在大多数情况下,它也是一种最优算法。

阈值估计有硬阈值估计和软阈值估计两种。硬阈值函数和软阈值函数分别如图 8.1(a)和(b)所示。硬阈值函数为

$$d_m(x) = h_T(x) = \begin{cases} x, & 若|x| > T \\ 0, & 若|x| \leqslant T \end{cases} \tag{8.2.14}$$

而软阈值函数为

$$d_m(x) = s_T(x) = \begin{cases} x - T, & 若 x > T \\ x + T, & 若 x < -T \\ 0, & 若|x| \leqslant T \end{cases} \tag{8.2.15}$$

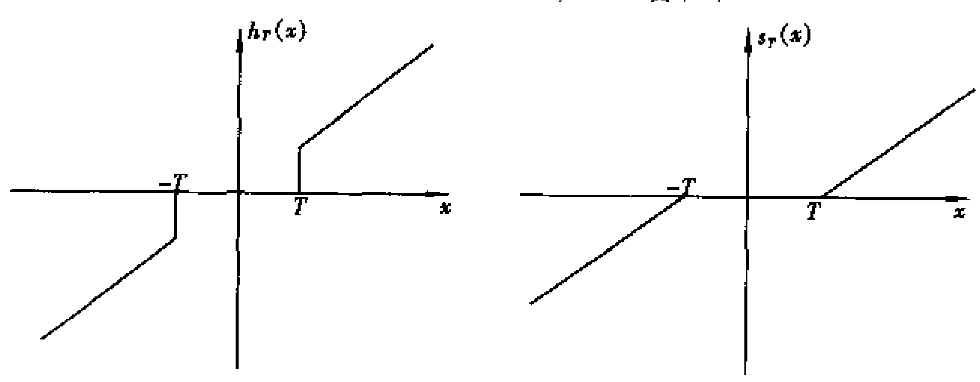


图 8.1 (a)硬阈值函数,(b)软阈值函数

显然,硬阈值函数有间断点,它只是简单的保留或者是去掉信号。而软阈值在进行阈值判别的同时也用阈值 T 对信号进行衰减。

对观测信号内积系数进行硬阈值判别表示为

$$d_m\big(X_B(m)\big) = h_T\big(X_B(m)\big) = \begin{cases} X_B(m), & 若|X_B(m)| > T \\ 0, & 若|X_B(m)| \leqslant T \end{cases} \tag{8.2.16}$$

代入(8.2.6)式得硬阈值估计风险

$$r_t(f) = \sum_{m=0}^{N-1} E\{|f_B(m) - h_T(X_B(m))|^2\} = \sum_{|X_B(m)|>T} \sigma^2 + \sum_{|X_B(m)|\leq T} |f_B(m)|^2 \tag{8.2.17}$$

因为 $X_B(m)=f_B(m)+W_B(m)$，所以当 $|X_B(m)|>T$ 时，并不能确保 $|f_B(m)|^2>\sigma^2$；同样，当 $|X_B(m)|\leq T$ 时，也不能确保 $|f_B(m)|^2<\sigma^2$，故上式中每一项不一定是 $\min[|f_B(m)|^2,\sigma^2]$，这意味着硬阈值的风险将大于 Oracle 投影估计，即

$$r_t(f) \geq r_p(f) \tag{8.2.18}$$

对观测信号内积系数进行软阈值判别表示为

$$d_m[x_B(m)] = s_T[X_B(m)] = \begin{cases} X_B(m) - T, & \text{若 } X_B(m) > T \\ X_B(m) + T, & \text{若 } X_B(m) < -T \\ 0, & \text{若 } |X_B(m)| \leq T \end{cases} \tag{8.2.19}$$

在得到最小风险 $r_{inf}(f)$ 时，如(8.2.7)式所示，用 $\alpha(m)$ 对 $X_B(m)$ 进行了衰减，从而抑制噪声。软阈值估计在保留 $X_B(m)$ 时也用阈值对它进行了衰减。

8.2.3　阈值风险

上面讨论硬阈值估计和软阈值估计时尚未涉及阈值如何确定的问题，Donoho 和 Johnstone 已证明：对适当选取的阈值，阈值估计的风险接近于 $r_p(f)$。用 O_d 表示所有对角算子的集合，N 表示信号长度，关于阈值估计风险有下述定理。

定理 8.3　设 $T=\sigma\sqrt{2\log_e N}$，对所有 $N\geq 4$，硬或软阈值估计风险 $r_t(f)$ 满足

$$r_t(f) \leq (2\log_e N + 1)\left(\sigma^2 + r_p(f)\right) \tag{8.2.20}$$

在所有 B 的对角估计算子中，因子 $2\log_e N$ 是最优的：

$$\lim_{N\to+\infty} \inf_{D\in O_d} \sup_{f\in C^N} \frac{E\{\|f-DX\|^2\}}{\sigma^2 + r_p(f)} \frac{1}{2\log_e N} = 1 \tag{8.2.21}$$

该定理说明阈值估计风险最多比 $r_p(f)$ 大 $2\log_e N$ 倍，且因子 $2\log_e N$ 不能由任何其他对角估计算子所改进。

还可以证明：用 $r_{inf}(f)$ 代替 $r_p(f)$ 时，上述定理仍然成立。

按上述定理选取的阈值 $T>\sigma$，即平均而言，阈值高于噪声电平，或者说阈值以极大的概率高于噪声电平，所以阈值与 σ 成正比是不难理解的。但阈值为什么会随样本数 N 的增加而增大呢？这是由于高斯分布的截断会随 N 的增加而产生越来越大的噪声系数。为什么 $r_t(f)$ 的上界会出现因子 $2\log_e N$ 呢？当 $|f_B(m)|<\sigma$ 时，由于 $T>\sigma$，$|X_B(m)|<T$ 的可能性很大，硬阈值估计和 Oracle 投影阈值估计都将 $X_B(m)$ 置零，不会造成 $r_t(f)$ 与 $r_p(f)$ 的很大差别；当 $|f_B(m)|\geq 2T$ 时，$|X_B(m)|\geq T$ 的可能性很大，硬阈值估计和 Oracle 投影阈值估计都将保留 $X_B(m)$，也不会造成 $r_t(f)$ 与 $r_p(f)$ 的很大差别。而当 $\sigma\leq|f_B(m)|<2T$ 时，Oracle 投影阈值估计产生的风险为 $\min[|f_B(m)^2|,\sigma^2]=\sigma^2$，如此时 $|X_B(m)|\leq T$，则硬阈值估计将 $X_B(m)$ 置零，产生的风险为

$$|f_B(m)|^2 \approx T^2 = 2\sigma^2\log_e N$$

也就是说，硬阈值估计的风险比 Oracle 投影阈值估计大 $2\log_e N$ 倍。一般来说，$|f_B(m)|$ 与 T 为同一数量级的部分还是很小的，所以阈值估计风险 $r_t(f)$ 比(8.2.20)式所示的上界小得多。在

这个意义下，阈值估计风险 $r_t(f)$ 接近理论极限 $r_{\inf}(f)$，是一种最优估计算法。

定理 8.3　假定噪声是白噪声，如噪声是有色噪声，则 $E\{|W_B(m)|^2\}=\sigma_m^2$。这时仍可用阈值估计算法去噪，且有如下定理：

定理 8.4　对 $0\leq m\leq N-1$，取 $T_m=\sigma_m\sqrt{2\log_e N}$，对所有 $N\geq 4$，硬或软阈值估计风险 $r_t(f)$ 满足

$$r_t(f)\leq(2\log_e N+1)[\bar{\sigma}^2+r_p(f)] \tag{8.2.22}$$

其中

$$\bar{\sigma}^2=\frac{1}{N}\sum_{m=0}^{N-1}\sigma_m^2 \tag{8.2.23}$$

8.2.4　Stein 无偏风险估计

Stein 无偏风险估计 SURE(Stein Unbiased Risk Estimator)，实际上是对上述阈值选取方法的一种改进。按定理 8.3 选取的阈值与信号的特性无关，SURE 阈值选取依赖于观测信号，具有自适应性。由于阈值选取与信号有关，故用 $\tilde{r}_t(f,T)$ 表示对风险的估计，使 $\tilde{r}_t(f,T)$ 最小的阈值记为 $\tilde{T}$。

当 $|X_B(m)|\leq T$ 时，软阈值估计置 $X_B(m)$ 为零，产生的风险为 $|f_B(m)|^2$，因为

$$E\{|X_B(m)|^2\}=|f_B(m)|^2+\sigma^2$$

所以可用 $|X_B(m)|^2-\sigma^2$ 来估计 $|f_B(m)|^2$；当 $|X_B(m)|>T$ 时，软阈值估计从 $X_B(m)$ 中减掉 T，产生的平均风险为

$$E\{[f_B(m)-X_B(m)+T]^2\}=E\{[T-W_B(m)]^2\}=T^2+\sigma^2$$

于是可以用观测值和噪声方差来估计风险 $r_t(f,T)$，即

$$\tilde{r}_t(f,T)=\sum_{m=0}^{N-1}\Phi(|X_B(m)|^2) \tag{8.2.24}$$

其中

$$\Phi(x)=\begin{cases}x-\sigma^2, & 若\ x\leq T^2\\ T^2+\sigma^2, & 若\ x>T^2\end{cases} \tag{8.2.25}$$

Donoho 和 Johnstone 已证明，$\tilde{r}_t(f,T)$ 是对软阈值估计风险 $r_t(f,T)$ 的无偏估计，即

$$E\{\tilde{r}_t(f,T)\}=r_t(f,T) \tag{8.2.26}$$

为了找到使 $\tilde{r}_t(f,T)$ 最小的阈值 $\tilde{T}$，将 $X_B(m)$ 按幅度大小递减排列，重新排列后得到的系数记为 $X_B^r(k)$。设 l 为某一指标，使得 $|X_B^r(l)|\geq T>|X_B^r(l+1)|$，于是

$$\tilde{r}(f,T)=\sum_{k=l}^{N-1}|X_B^r(k)|^2-(N-l)\sigma^2+l(T^2+\sigma^2) \tag{8.2.27}$$

对所有的 $0\leq l\leq N-1$，逐一按上式计算 $\tilde{r}_t(f,T)$，就可以找到最小的 $\tilde{r}_t(f,T)$ 和相应的阈值 $\tilde{T}$。

虽然 $\tilde{r}_t(f,T)$ 是 $r_t(f,T)$ 的无偏估计，但其方差可能导致产生太小阈值 $\tilde{T}$ 的差错。当信噪比很小，即 $\|f\|^2\ll E\{\|W\|^2\}=N\sigma^2$ 时就会出现这种情况。此时应该用 $T=\sigma\sqrt{2\log_e N}$ 作为阈值以滤除噪声。由于 $E\{\|X\|^2\}=\|f\|^2+N\sigma^2$，所以可用 $\|X\|^2-N\sigma^2$ 来估计 $\|f\|^2$ 并将它与最小

能量级 $\varepsilon_N = \sigma^2 N^{1/2}(\log_e N)^{3/2}$ 比较，所得 SURE 阈值如下：

$$T = \begin{cases} \sigma\sqrt{2\log_e N}, & 若\|X\| - N\sigma^2 \leqslant \varepsilon_N \\ \tilde{T}, & 若\|X\| - N\sigma^2 > \varepsilon_N \end{cases} \tag{8.2.28}$$

小波去噪涉及两个问题：如何选取阈值和怎样实现取阈值（即取阈值的方式）。Matlab 中有 4 种选取阈值的规则，其中第 1 种'sqtwolog'就是按定理 8.3 确定阈值；第 2 种'rigrsure'是按 SURE 方法确定阈值；第 3 种'heursure'实际上是前两种的混合，如(8.2.28)式所示，当信噪比较大时，按第 2 种规则选取阈值，当信噪比较小时，按第 1 种规则选取阈值；第 4 种是'minmax'。函数 thselect 用来计算相应规则的阈值。取阈值的方式有硬阈值方式和软阈值方式两种。

8.3　小波去噪

对很大一类信号而言，它们的能量高度集中在少量幅度较大的小波系数或小波包系数中，而噪声由于其随机性，能量却比较均匀地分布于变换域，所以可以在变换域用简单的阈值判别法在保留信号大幅值分解系数的同时最大限度地抑制噪声。

设信号在某一尺度 2^L 上的离散逼近 $f(n)$ 被加性噪声 $W(n)$ 污染，观测数据 $X(n) = f(n) + W(n)$。将 $X(n)$ 在正交规范基

$$B = [\{\phi_{J,m}(n)\}_{m\in Z}, \{\psi_{j,m}(n)\}_{L<j\leqslant J, m\in Z}]$$

上分解，小波去噪是对分解系数取阈值后进行重构，即对 f 的估计可写成

$$\tilde{F} = \sum_{j=L+1}^{J}\sum_{m}\rho_T(\langle X, \psi_{j,m}\rangle)\psi_{j,m} + \sum_{m}\rho_T(\langle X, \phi_{J,m}\rangle)\phi_{J,m}$$

其中，ρ_T 表示对分解系数取硬阈值或软阈值。一般来说，对小波系数的阈值电平与逼近系数的阈值电平不同，往往是完全保留逼近系数而只对小波系数取阈值，而且在不同尺度上的阈值电平也可以不一样。对比较平滑和分段正则信号，信号的能量主要集中在低频部分，所以应保留逼近系数，而噪声能量主要集中在高频部分，所以应对小波系数取阈值，由于阈值电平以很大概率高于噪声电平，将低于阈值电平的小波系数置零就在很大程度上滤除了噪声，信号的锐变部分会产生大幅值的小波系数，保留这些大幅值的小波系数避免了对信号锐变部分的平滑，从而保留了这些有重要特征的信号细节，这正是小波去噪与基于傅立叶变换的频域滤波方法的根本区别。由于傅立叶变换不具有时域局部性，所以在滤除噪声的同时也平滑了信号的锐变部分，而小波变换具有时域局部性，所以小波去噪相当于一个自适应平滑过程，它只在信号的正则部分平滑掉噪声，而在其锐变部分保留了信号的细节。也可以说，小波去噪是一个幅值域滤波的过程。

在用阈值判别法从被噪声污染的观测数据中估计信号时，必须选择适当的阈值电平，要确定阈值，又必须知道噪声方差。当我们对噪声的先验知识不足，不知道噪声方差的情况时，可以通过小波分解来估计噪声方差。当信号是分段正则的函数时，用最精细尺度上的小波系数的中值来计算噪声方差是一个强健估计。具体来说，如 M 是 N 个独立的零均值，方差为 σ^2 的高斯随机变量绝对值的中值，则不难证明

$$E\{M\} \approx 0.6457\sigma$$

如信号比较平滑，对观测数据作小波分解之后，信号能量主要包含在逼近系数中，由于噪声能量主要分布在高频区域，所以它主要集中在最精细尺度上的小波系数中。当信号分段正则时，如小波有足够高的消失矩，只有信号的少数孤立锐变点产生少量大幅值小波系数，所以对小波系数的中值影响很微小。当小波支集不包含两个以上信号锐变点时，可以认为这些锐变点是彼此孤立的，孤立锐变点在不同尺度上产生相同数量的大幅值小波系数。如有 N 个观测数据，最精细尺度上的小波系数有 $N/2$ 个，其他尺度上的小波系数都少于 $N/2$，所以信号孤立锐变点对最精细尺度上的小波系数的中值影响最小，小波系数中值是对噪声方差的强健估计。设 M_X 是 $X(n)$ 在最精细尺度上的小波系数绝对值的中值，则对零均值，方差为 σ^2 的高斯白噪声的方差估计为

$$\hat{\sigma} = \frac{M_X}{0.6457} \tag{8.3.1}$$

如原始信号是一个有色零均值平稳高斯序列，那么分解系数仍然是高斯分布的，但对每个分解级次而言，系数是有色零均值平稳高斯序列。如定理 8.4 所指出的那样，不同级次的高斯序列有不同的方差，但仍然可以用类似方法去估计噪声方差。在用阈值方法去噪时每一级应该选取不同的阈值。Matlab 函数 wdencmp 可以用来实现噪声消除，其中有 1 个可选择的字符串变量，'gbl'为所有级次都选择相同的阈值，'lvd'则对不同级次选择不同阈值。

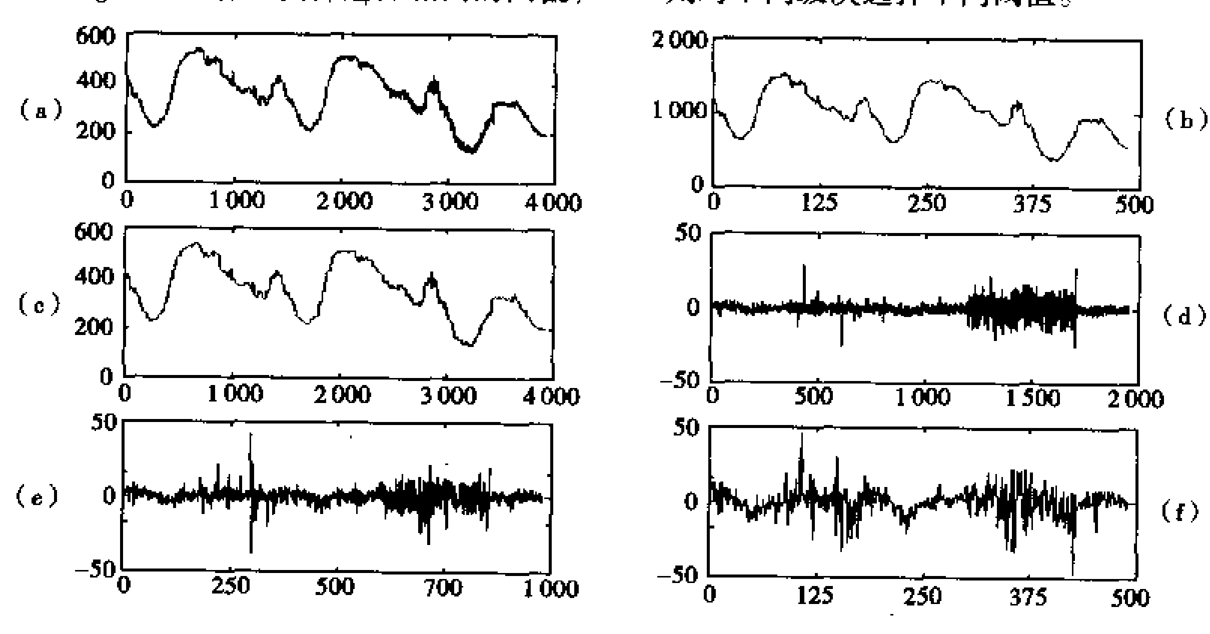


图 8.2 小波去噪

例 8.2 图 8.2(a) 为 3 天电力消耗情况曲线，由于监测设备的缺陷导致监测噪声，用小波分析方法可以有效地去除噪声。小波去噪分为 3 步：第 1 步进行小波分解，做 3 级分解后得到的逼近系数为图 8.2(b)，第 1、2、3 级的细节系数分别示于图 8.2(d)、(e)、(f)。一般来说，噪声能量主要集中在第 1 级细节系数，由图 8.2(d) 可以清楚地看到在最后一段时间出现了更明显的噪声干扰。相对于快速起伏的噪声而言，信号变化比较平缓，它主要保留在逼近系数之中。第 2 步对小波系数取阈值，最粗略的办法就是将所有的细节系数置 0，但这样在抑制噪声的同时也平滑了信号的快速变化。小波幅值域滤波方法的优越性就在于有效地抑制噪声的同时也保留了信号的快速变化，所以对细节系数取软阈值（阈值为 9.2582），只保留那些幅值较大的细节系数。第 3 步为信号重构。8.2(c) 为取软阈值后重构的信号，可以明显地看到在有

效抑制噪声的同时确实也保留了信号的快速变化。

例 8.3　图像信号去噪与一维情况基本一样，但在用定理 8.3 计算阈值时，信号长度应该代之以图像像素数。图 8.3 显示了小波图像去噪的情况，与图 5.5 的原始图像比较，在获得可以接收的去噪效果的同时，图像的边缘和纹理结构也变得比较模糊了。



噪声图像



去噪图像

图 8.3　小波图像去噪

8.4　数据压缩的信息论基础简介

数据压缩编码系统如图 8.4 所示。按照信息论的观点，信号源（简称信源）是由一系列符号构成的，记为 $A=\{x_k\}_{1\leq k\leq K}$。对一维信号而言，一系列符号实际上就是一系列离散样本，离散样本值的大小用二进制数字表示，称为码字。信源符号经过脉冲编码调制（PCM）以后，变成二进制码流。PCM 的作用主要是使信息得以在数字信道上传输。压缩编码器对信源符号进一步编码，达到数据压缩的目的，以便于存储或提高信息传输的比特率。信息在数字信道上传输时会受到各种不同类型噪声的干扰，信道编码器的作用主要是减小误码率，提高系统的抗干扰能力。下面着重讨论信源编码问题。

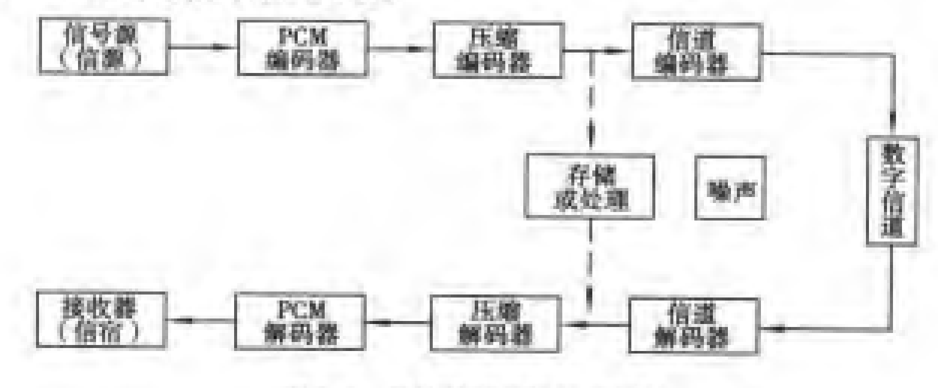


图 8.4　压缩编码系统原理框图

8.4.1　信息量与熵

在信息传输系统中，对信息接收方（称为信宿）而言，在未收到信源发出的符号之前，它不

确定信源到底会发出什么符号，以 p_k 表示符号 x_k 出现的概率，它度量了这种不确定性的程度。在理想情况下，信息传输和接收的过程中没有噪声的干扰，信宿接收到符号之后，这种不确定性就完全消除了，即使有噪声干扰，也会在一定程度上消除不确定性。接收信息的过程实际上就是消除这种不确定性的过程，信息量度量了消除不确定性的程度。某一符号 x_k 出现的概率 p_k 越小，接收到它之后不确定性消除的程度越高，所以可用 $-\log_2 p_k$ 来度量其信息量，则信源的平均信息量为

$$H = -\sum_{k=1}^{K} p_k \log_2 p_k \tag{8.4.1}$$

其单位为比特(bit)，H 又称信息熵。这样定义的熵没有考虑信源符号之间的关联，称为无记忆信源熵，或0阶熵，记为 H_0。

在信源编码中，用二进制数字来表示信源符号，称为码字，设 l_k 为 x_k 的码字长度，则平均码字长度为

$$R = \sum_{k=1}^{K} p_k l_k \tag{8.4.2}$$

根据信息论中信源编码理论，信息熵是平均码字长度的理论极限，即 $R \geqslant H$；而且存在无失真的编码方法，使得 $R \leqslant H+1$。

假定数字灰度图像的灰度级为 2^n，这时不同的信源符号实际上就是不同的灰度值，即 $K=2^n$。假设所有灰度级出现的概率相等，均为 $K^{-1}=2^{-n}$，则信息熵 $H=n$。显然，对有 $K=2^n$ 个灰度级的图像，且所有灰度级等概率出现时，至少要用平均码字长度为 n 比特的码字编码。如将灰度级的高低与码字的值对应起来，就是 PCD 码。当某一灰度等级 x_i 出现的概率为 1 时，其他灰度等级出现的概率必然为 0，这是另一种极端情况，此时信息熵等于零。这时信宿接收的信源符号是完全确定的，这样的信源完全没有任何信息量。

在上面的讨论中，假定任一灰度级出现的概率与其他灰度级无关。实际上，同一帧图像相邻像素间，以及运动图像相邻帧之间的灰度级是有很强的相关性的。假定信源符号 x_p 出现后又出现 x_k 的条件概率为 $p(x_k|x_p)$，符号 x_k 和 x_p 不论前后顺序同时出现的概率为 $p(x_k,x_p)$，则定义条件熵

$$H(x_k \mid x_p) = -\sum_{k=1}^{K}\sum_{p=1}^{K} p(x_k,x_p)\log_2 p(x_k \mid x_p) \tag{8.4.3}$$

这样定义的熵称为1阶熵，记为 H_1。依次类推还可以定义2阶和3阶等高阶熵，而且可以证明

$$H_0 \geqslant H_1 \geqslant H_2 \geqslant H_3 \geqslant \cdots \tag{8.4.4}$$

当 x_k 和 x_p 独立，即 $p(x_k,x_p)=p(x_k)p(x_p)$ 时，$H_1=H_0$；而当 x_k 和 x_p 有一定程度的相关性时，$H_0>H_1$，这是因为接收到 x_p 后，是否再接收到 x_k 的不确定性减小了。信源符号之间有相关性的信源，其高阶熵小于0阶熵，存在信息冗余。当信源存在信息冗余时，可以通过去除冗余，达到压缩平均码字长度，提高编码效率的目的，这是数据压缩的主要理论依据。对语音信号和图像信号通过适当选择的正交基进行变换，变换域内系数之间的相关性大大减少，再对变换系数进行编码就可以达到数据压缩的目的，这就是变换编码的理论依据。在小波变换兴起之前，图像压缩主要利用离散余弦变换(DCT)，例如静止图像的国际压缩标准 JPEG 以及运动图像的国际压缩标准 MPEG—1 和 MPEG—2 都采用了 DCT，而在 MPEG—2000 标准中，则采用了小

波变换。例如灰度级为 2^n 的图像，无论它的像素之间有多么强的相关性，如直接对灰度编码，平均码字长度至少为 n 比特。如对原始图像作多尺度分解，细节图像中只有少量系数幅度较大，将某一阈值以下的系数置 0，就可达到数据压缩的目的。

8.4.2　压缩比、编码效率与率失真函数

变换编码仅仅是诸多压缩编码方法之一。压缩比与编码效率是衡量编解码设备性能的重要指标。

设原始信号的平均码长为 $\overline{R}$，压缩后的平均码长为 $\overline{R}_c$，则定义压缩比

$$C = \frac{\overline{R}}{\overline{R}_c} \tag{8.4.5}$$

在一般情况下，编码效率往往用下列简单公式表示

$$\eta = \frac{H}{R} \tag{8.4.6}$$

其中，R 为平均码长，它既可以是 $\overline{R}$，也可以是 $\overline{R}_c$。

如已编码信号经解码后可以完全恢复原始信号，称为无失真编码；否则为有失真编码。根据信息论中信源编码理论，可以证明在 $R \geqslant H$ 的条件下，总可以设计出某种无失真编码方法。如编码结果使 R 远大于 H，表明这种编码方法效率很低，占用比特数太多。例如对图像样本量化值直接采用 PCD 编码，则平均码长 R 就远大于图像熵 H。

在信息论中，其率失真理论回答了有关失真度与信道容量之间的关系问题，为限失真编码奠定了理论基础。率失真函数定义如下：

定义 8.2　在传输中使信号的失真度小于或等于某一值 D 所必须的信道容量的最小值 $R(D)$ 称为信息速率-失真函数，或率失真函数。

在给定失真度条件下，率失真函数定义了信源编码所能达到的比特率下限。率失真函数和失真的度量标准与信源的统计特性有着密切的关系。在图像处理与图像编码技术中，通常用均方误差作为图像失真的度量标准。根据信息论的结论，在假定图像信源是正态分布的条件下，率失真函数为

$$R(D) = \begin{cases} \dfrac{1}{2}\log_2 \dfrac{\sigma^2}{D}, & 0 \leqslant D \leqslant \sigma^2 \\ 0, & D > \sigma^2 \end{cases} \tag{8.4.7}$$

其中，D 为允许的均方误差，σ^2 为信号的方差。在图像编码中，$R(D)$ 表示在均方误差不大于 D 时传输一个像素的平均码长。上式还指出：如 $D > \sigma^2$，即允许的失真比信号的方差还大，则 $R(D) = 0$，信息传输已失去意义。另外由上式可见，所需传输的信号的方差越小，允许的失真越大，所需传输的比特率越低。

还可以证明：如果图像信号是非正态分布的，则 $R(D)$ 的数值比（8.4.7）式给出的数值更低，因为在一定功率条件下，正态分布时信源的熵最大。此时（8.4.7）式的意义在于，在给定信号功率和允许失真度的条件下，它给出了信源编码可以达到的比特率下限。率失真理论虽然没有给出怎样达到比特率下限的具体方法，但从理论上指明了方向。即在给定允许失真度的条件下，为了减少传输的比特率，应尽量减小信号的方差。

对于绝大多数图像来说，在局部空间和时间上是高度相关的，因而可以在已接受到的像素

的基础上对当前像素进行预测。由条件熵的性质可知，在已知 $x_{k-1},x_{k-2},\cdots,x_{k-m}$ 的条件下来确定 x_k 的熵，其值总是比较小的。记 x_k 的预测值为 $\hat{x}_k$，它很接近 x_k，所以 $d_k=x_k-\hat{x}_k$ 的方差 σ_d^2 比原始图像的方差 σ_x^2 肯定要小。因此，传输 d_k 就可用较少的比特数，这就是所谓的预测编码，它在图像压缩编码中获得了广泛应用。

8.4.3　熵编码

熵编码的目的是使编码后的平均码字长度尽可能接近信息熵。一般是根据信源符号出现的概率赋予不同长度的码字，概率大的符号用短码字，概率小的符号用长码字，由(8.4.2)式可以看到，这样获得的平均码字较短。这种编码方法称为变字长编码。

定理 8.5　变长最佳编码定理。

在变字长编码中，对出现概率大的符号赋予短码字，而对出现概率小的符号赋予长码字。如码字长度严格按照所对应符号出现概率的大小逆序排列，则编码结果平均码字长度一定小于任何其他排列方式。

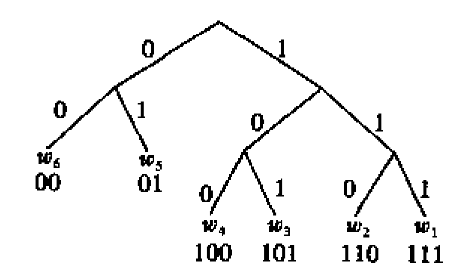


图 8.5　惟一可译编码

霍夫曼编码就是一种最佳编码算法。在介绍霍夫曼编码方法之前，先简要叙述惟一可译编码。某些情况下，为了减少平均码字长度，往往在码字之间不加同步码。这就要求所编码字序列能被惟一地译出来，满足这个条件的编码称为惟一可译编码，也常称为单义可译码。设想信源有 4 个符号 $\{x_k\}_{1\leq k\leq 4}$，其相应的码字为

$$\{w_1=0,w_2=10,w_3=110,w_4=101\}$$

则码字序列 1010 既可译成 w_2w_2，又可译成 w_4w_1，所以上述编码不是惟一可译编码。为得到惟一可译编码，必须保证任一码字都不是另一码字的前缀（或起始），在上述编码中，w_2 是 w_4 的前缀，所以它不是惟一可译编码。而下列编码

$$\{w_1=0,w_2=10,w_3=110,w_4=111\}$$

就是惟一可译编码。惟一可译编码可表示为二叉树，例如图 8.5 表示有 6 个信源符号 $\{x_k\}_{1\leq k\leq 6}$，如某一码字已对应二叉树的一个节点，该节点就不能再往下分枝，因为父节点必然是子节点的前缀。

定理 8.6　霍夫曼编码。

设信源符号集为 $\{x_k\}_{1\leq k\leq K}$，按其出现概率的大小排列如下

$$\{(x_1,p_1),(x_2,p_2),(x_3,p_3),\cdots,(x_k,p_k)\} \tag{8.4.8}$$

其中，$p_k\leq p_{k+1}$。将具有最小概率的两个符号 x_1 和 x_2 结合成一个符号 $x_{1,2}$，其概率为

$$p_{1,2}=p_1+p_2$$

从而得到有 $K-1$ 个符号的序列

$$\{(x_{1,2},p_{1,2}),(x_3,p_3),\cdots,(x_k,p_k)\} \tag{8.4.9}$$

再将上述序列中具有最小概率的两个符号结合成一个符号。过程中每结合一次就从下往上由两个子节点生成一个父节点，直到形成一个完整的二叉树。从根节点开始，左右两个分枝分别标以 0 和 1，就得到霍夫曼编码。

例 8.4　设信源有 6 个信源符号 $\{x_k\}_{1\leq k\leq 6}$，它们的出现概率依次为

$$\{p_1=0.05,p_2=0.1,p_3=0.1,p_4=0.15,p_5=0.2,p_6=0.4\}$$

用霍夫曼编码生成的二叉树及码字如图 8.6 所示。由符号 x_1 和 x_2 结合成一个符号 $x_{1,2}$，它的概率为 $p_{1,2}=p_1+p_2=0.15$。第二次迭代过程中，最低概率的两个符号是 x_3 和 $x_{1,2}$，将它们结合为 $x_{1,2,3}$，其概率为 0.25。第三次迭代过程中，最低概率的两个符号是 x_4 和 x_5，将它们结合为 $x_{4,5}$，其概率为 0.35。第四次迭代过程中，最低概率的两个符号是 $x_{1,2,3}$ 和 $x_{4,5}$，将它们结合为 $x_{1,2,3,4,5}$，其概率为 0.6。最后将 $x_{1,2,3,4,5}$ 和 x_6 结合到达根节点，概率为 1。生成的码字为

$$\{1111,1110,110,101,100,0\}$$

平均码字长度为 2.35。信源熵为 2.28，霍夫曼编码的平均码字长度已很接近信源熵。

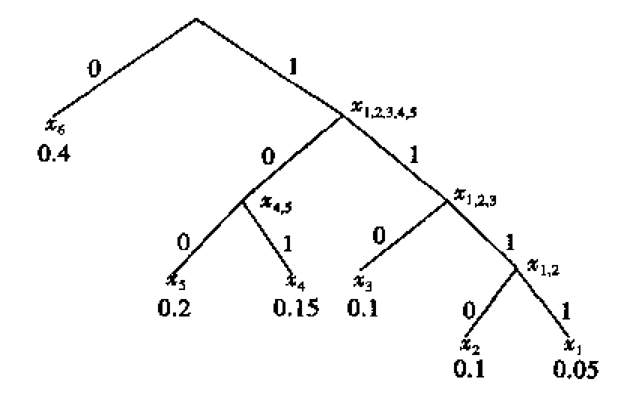


图 8.6　霍夫曼编码

除霍夫曼编码之外，还有其他熵编码方法，如算术编码和游程编码。它们都是无失真的编码方法。而下面将要介绍的变换编码是一种有失真编码。

8.5　变换编码

变换编码的基本思想是去相关。变换、量化和熵编码是编码器的 3 个主要组成部分，如图 8.7 所示。变换的目的是去相关，通过选择适当的基，使信号在变换域内的系数尽可能不相关。变换系数是无限精度的实数，但只能用有限字长的二进制数表示，这一过程称为量化。显然，量化必然会产生误差，所以变换编码是有失真编码，量化误差是图像编码的主要误差。量化器的设计要在给定量化字长的条件下，得到最小量化误差；或者在给定量化误差的前提下，设计具有最短平均字长的量化器。量化器输出的量化系数最后再进行熵编码。熵编码问题已在上一节做了简单介绍，本节主要讨论基的选择和量化器设计这两个问题。

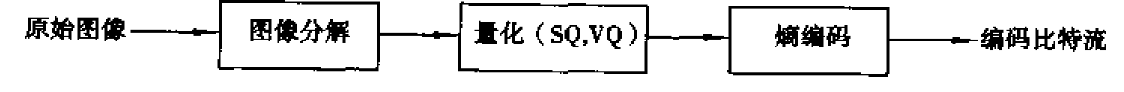


图 8.7　图像变换编码过程

将信号视为随机过程,变换编码器在正交规范基 $B=\{g_m\}_{0\leq m\leq N-1}$ 下分解信号 X,

$$X=\sum_{m=0}^{N-1}X_B(m)g_m \tag{8.5.1}$$

其中,每一变换系数

$$X_B(m)=\langle X,g_m\rangle=\sum_{n=0}^{N-1}X(n)g_m^*(n) \tag{8.5.2}$$

都应视为随机变量,也就是一个信源符号。$X_B(m)$可取任意实数值,为构造有限字长码,每一系数 $X_B(m)$由其量化值 $\tilde{X}_B(m)$近似,$\tilde{X}_B(m)$只能在有限实数集合上取值。一般应对量化系数再进行熵编码,例如前面提及的对灰度图像的熵编码。从量化系数重构的信号是

$$\tilde{X}=\sum_{m=0}^{N-1}\tilde{X}_B(m)g_m \tag{8.5.3}$$

显然,$\tilde{X}$ 仅仅是信号 X 的逼近,所以变换编码是有失真编码。定义失真率

$$d=E\{\|X-\tilde{X}\|^2\}=\sum_{m=0}^{N-1}E\{|X_B(m)-\tilde{X}_B(m)|^2\} \tag{8.5.4}$$

来度量 $\tilde{X}$ 对信号 X 的逼近程度。不难理解,量化器的平均码字长度越长,失真率越小,但平均码字长度越长,量化器的编码效率越低,压缩比越小。所以必须在失真率和编码效率之间进行折中,量化器的设计应该在给定平均码字长度的条件下使(8.5.4)式所示的均方误差最小,或在给定失真率的条件下使平均码字长度最短。

8.5.1 量化器

量化器有标量量化和矢量量化两种方式。标量量化独立地逼近每一个内积系数 $X_B(m)$。通过选择适当的正交基,使变换域上的内积系数相关性很小时,矢量量化对量化器性能的改进很少,而且标量量化的计算量较少,所以标量量化较常使用。以下主要讨论标量量化器。

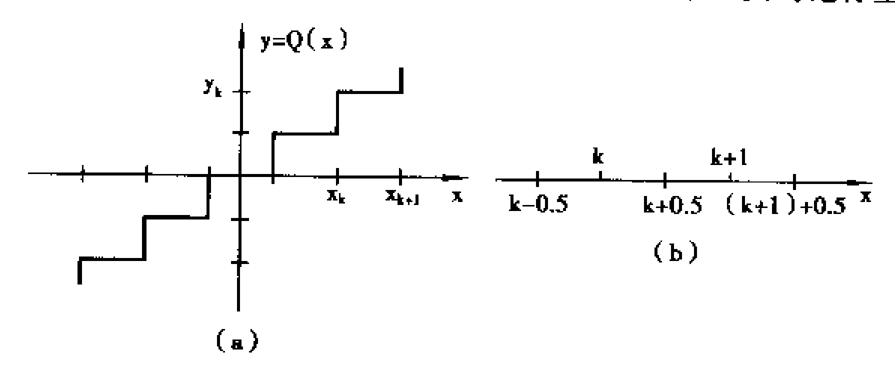


图 8.8 标量量化器

标量量化器的输出-输入关系 $y=Q(x)$如图 8.8 所示,输入 x 为连续实数,输出 y 为不连续的(即量化的)实数。设 x 在$[a,b]$区间连续取值,将$[a,b]$划分为 K 个长度不一定相等的量化区间$\{(x_{k-1},x_k)\}_{1\leq k\leq K}$,其中,$x_0=a,x_k=b$。当 $x\in(x_k,x_{k+1})$时,标量量化器将属于这一量化区间的 x 值都用一个确定的值 y_k 来逼近,即 $Q(x)=y_k$,如图 8.8(a)所示。对变换编码而言,量化器的输入 $X_B(m)$是连续取值的,输出 $\tilde{X}_B(m)$的值是量化的。四舍五入就是量化器的一个简单例子,这时将整个实轴划分为等长的量化区间$[k-0.5,k+0.5]_{k\subset Z}$,当 $x\in[k-0.5,k+0.5]$

时，$Q(x)=k$，如图 8.8(b)所示。

假设随机变量 X 的概率密度函数 $p(x)$ 已知，则量化均方误差

$$d = E\{(X-\tilde{X})^2\} = \int_{-\infty}^{+\infty} \left(x - Q(x)\right)^2 p(x)\mathrm{d}x \tag{8.5.5}$$

若量化区间的长度 $\Delta_k = x_k - x_{k-1}$ 足够小，则 $X \in (x_{k-1}, x_k)$ 的概率为

$$p_k = p(x)\Delta_k \tag{8.5.6}$$

很容易证明[3]：

定理 8.7　当 $x \in (x_k, x_{k+1})$ 时，如 $Q(x) = (x_k + x_{k-1})/2$，则量化均方误差最小，且

$$d = \frac{1}{12}\sum_{k=1}^{K} p_k \Delta_k^2 \tag{8.5.7}$$

上述定理意味着，当量化区间的长度足够小时，使量化误差最小的量化器应取量化值为量化区间的重心。

均匀量化器是很重要的一种量化器，这时所有量化区间的长度相同，均为

$$\Delta = x_k - x_{k-1}, 1 \leqslant k \leqslant K \tag{8.5.8}$$

定理 8.7 告诉我们，如 Δ 足够小，最小量化均方误差

$$d = \frac{\Delta^2}{12}\sum_{k=1}^{K} p_k = \frac{\Delta^2}{12} \tag{8.5.9}$$

与信源的概率密度函数无关。

上面讨论了使量化误差最小的量化器，下面讨论在给定量化误差的前提下，如何设计量化器使平均码字长度最短。熵编码理论告诉我们，信息熵规定了平均码字长度的理论极限。设量化器为 $\tilde{X} = Q(X)$，根据信息熵的定义

$$H(\tilde{X}) = -\sum_{k=1}^{K} p_k \log_2 p_k \tag{8.5.10}$$

由于 X 是连续取值的，定义所谓的微分熵如下

$$H_d(X) = -\int_{-\infty}^{+\infty} p(x)\log_2 p(x)\mathrm{d}x \tag{8.5.11}$$

如量化区间的长度足够小，则可以证明[3]：

$$H(\tilde{X}) \geqslant H_d(X) - \frac{1}{2}\log_2(12d) \tag{8.5.12}$$

当且仅当量化器为均匀量化器时，上式中等号成立。

上述结论意味着均匀量化器使得平均码字长度最短，且最短平均码字长度为

$$R_X = H(\tilde{X}) = H_d(X) - \log_2\Delta \tag{8.5.13}$$

此时量化误差为

$$d(R_X) = \frac{1}{12}2^{2H_d(X)}2^{-2R_X} \tag{8.5.14}$$

8.5.2　比特分配

信号 X 经过正交变换之后，量化器将每一个内积系数 $X_B(m)$ 量化为 $\tilde{X}_B(m)$。由上面讨

论知道,采用均匀量化器将使得量化误差 $d_m = E\{|X_B(m) - \tilde{X}_B(m)|^2\}$ 最小,且平均码字长度 $R_m = H_d\big(X_B(m)\big) - \log_2\Delta_m$,其中 Δ_m 是对 $X_B(m)$ 的量化区间长度。现在的问题是如何优化选择 $\{\Delta_m\}_{0\leqslant m\leqslant N-1}$,对给定的总量化误差

$$d = \sum_{m=0}^{N-1} d_m \tag{8.5.15}$$

使得总平均码字长度

$$R = \sum_{m=0}^{N-1} R_m \tag{8.5.16}$$

最短。

以 $\overline{R} = R/N$ 表示每个样本的平均码字长度,则可以证明[3]:

定理 8.8 如对所有内积系数,量化区间长度都相等,且

$$\Delta_m^2 = \frac{12d}{N}, 0 \leqslant m \leqslant N - 1 \tag{8.5.17}$$

以及对给定的总量化误差

$$d(\overline{R}) = \frac{N}{12}2^{2\overline{H}_d}2^{-2\overline{R}} \tag{8.5.18}$$

其中,

$$\overline{H}_d = \frac{1}{N}\sum_{m=0}^{N-1} H_d[X_B(m)] \tag{8.5.19}$$

为平均微分熵,这时总平均码字长度最短。

上述定理意味着,如沿基 B 的每个方向 g_m 都设置相等的量化区间长度从而引入相等的均方误差 $d_m = \Delta_m^2/12 = d/N$,那么量化器是最优的。这时用来对 $X_B(m)$ 进行量化编码的平均码字长度

$$R_m = H_d[X_B(m)] - \frac{1}{2}\log_2\left(\frac{12d}{N}\right) \tag{8.5.20}$$

最短,且它仅取决于 $X_B(m)$ 的微分熵。

8.5.3 最优基

从定理 8.8 可以看到,为使总平均码字长度最短,应选择适当的基使平均微分熵最小。一般说来,$X_B(m) = \langle X, g_m\rangle$ 的概率密度函数与各基矢量 g_m 的依赖关系非常复杂,难以计算。但当信号是高斯随机过程时,变换域系数 $X_B(m)$ 在任何基下都是高斯随机变量,其概率密度函数仅取决于方差 σ_m^2

$$p_m(x) = \frac{1}{\sqrt{2\pi}\sigma_m}\exp\left(-\frac{x^2}{2\sigma_m^2}\right)$$

从而可算出 $X_B(m)$ 的微分熵

$$H_d[X_B(m)] = \log_2\sigma_m + \log_2\sqrt{2\pi e}$$

代入(8.5.18)式得总量化误差

$$d(\overline{R}) = \frac{\pi e}{6}N\rho^2 2^{-2\overline{R}} \tag{8.5.21}$$

其中，ρ^2 是所有方差的几何平均

$$\rho^2 = \left(\prod_{m=0}^{N-1}\sigma_m^2\right)^{1/N} \tag{8.5.22}$$

这意味着最优基应使 ρ^2 最小。

可以证明：在信号 X 的 K-L 基下，几何平均方差 ρ^2 最小，从而总量化误差最小。当信号可以用高斯随机过程模型描述时，$\{X_B(m)\}_{1\leq m\leq N-1}$ 是联合高斯分布的，而在 K-L 基的分解下，它们又是互不相关的，所以它们是相互独立的。分解系数的独立性说明用标量量化而不用矢量量化的合理性。

8.6　小波变换编码

最新的小波编码器都沿袭变换编码的基本思想，即去相关性。同样的，小波编码器也由变换、量化和熵编码等 3 个主要部分构成。目前，人们围绕这 3 个方面的内容开展小波编码器的研究。将原始图像经二维小波变换后，转换成小波域上的小波系数，然后对小波系数进行量化编码。由于小波变换后使得原始图像能量集中在少量小波系数上，因此最简单的系数量化方法就是将某一阈值以下的系数滤去（或者表现为恒定常数），只保留那些能量较大的小波系数进行熵编码，从而达到数据压缩的目的。在这里，所采用的标量量化方法是分别在不同分辨率（不同分解层次）的小波频带来完成的。

8.6.1　子带编码

小波编码在实现机构和方法上与子带编码非常一致，可以说小波编码是子带编码的特例。子带变换器是一个多通道数字信号处理系统，如图 8.9 所示。显然，图 5.1 所示双通道多采样率滤波器组实际上是子带变换器的一种特殊情况。图 8.9 中的左边是子带变换的正变换，$H_0, H_1, \cdots, H_{M-1}$ 称为分析滤波器组，对信号进行滤波之后，输出的信号被抽取，即每 M 个离散值只保留一个。图 8.9 的右边是子带变换的逆变换，被滤波和抽取的信号首先经过插零处理，接着通过综合滤波器组 $G_0, G_1, \cdots, G_{M-1}$ 最后重构出信号。当分析滤波器组和综合滤波器组满足一定条件时，原始信号能从这些 M 组抽取的数值中完全重建，这些条件称为理想重构条件。一般来说，子带编码并不强求使用理想重构条件，但用小波编码时则使用严格的理想重构滤波器，并且要求符合正则性条件。

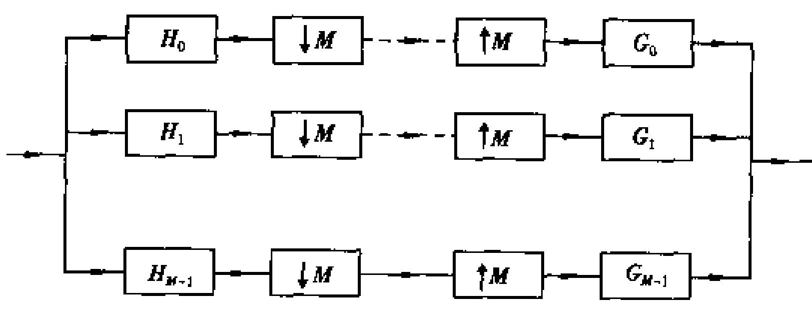


图 8.9　子带变换

子带编码的基本思想是，将原始信号通过一组滤波器滤波后并抽取，被分成若干个包含有

限频宽的子带,然后对每一个子带根据其不同的频率分布特性采取不同的编码方案,如标量量化或者矢量量化等等。

子带编码的优点在于:①由于每个子带都独立处理,因此其编码和解码的误差不会扩散到其他子带;②通过改变不同子带内的比特分配,整体的编码误差以视觉最优的方式分布在各个子带中。产生子带的过程本身并没有压缩信号,各个子带里的采样点数之和仍然等于原始信号中的采样点数,但是原始信号的能量在各个子带得到重新分配,使得独立对子带系数进行编码比直接对原始信号编码更为有效。

二维小波变换实际上是用一组相互独立的二维小波基函数将一幅图像分解成 4 部分,也就是将一幅图像分解成 4 个子带,每个子带包含着原始图像的不同频率成分,而且具有一定的空间取向。

8.6.2 小波基及分解级数的选取

虽然选择最优小波基用于图像编码是一个非常棘手的问题,但还是存在一些设计标准,如平滑性、逼近精度、支撑大小和滤波频率选择等,这些都是很重要的特性。如何最佳地组合这些特性仍然有待于人们的深入研究。

正交小波基的平滑性与消失矩对图像压缩效果有一定的影响,并且平滑性的影响要高于消失矩。例如 Harr 小波基由于不连续,会造成恢复图像中出现方块效应,而采取其他光滑的小波基则会消除方块效应。在实际应用中,一般要选择具有平滑特性的小波基,同时为了简化计算,只考虑一阶和二阶导数连续的小波基。在逼近理论框架下,样条小波非常适合用于图像压缩编码。

由于小波变换过程实际上是信号与滤波器卷积的过程,因此,滤波器的长度增加将导致卷积运算量增加;并且从边界延拓来看,滤波器长度越长,延拓的点越多,造成图像恢复的失真也越大,应适中地选择滤波器长度。目前,人们普遍选择 B 样条小波基所对应的双正交滤波器用于图像压缩编码,例如 9/7 双正交滤波器(9 和 7 分别为分解和重构低通滤波器长度)。此外,双正交小波基所构成的滤波器的相位是线性的。这也是人们常选择双正交小波基的理由。9/7 双正交滤波器构成的尺度函数和小波如图 5.4 所示,其滤波器系数为:

$h =$ 0.037 8 −0.023 8 −0.110 6 0.377 4 0.852 7 0.377 4 −0.110 6 −0.023 8 0.037 8

$\tilde{h} =$ −0.064 5 −0.040 7 0.418 1 0.788 5 0.418 1 −0.040 7 −0.064 5

虽然小波分解/重构的级数越多,产生的子带越多,对频带的划分越细,有利于编码;但是由于上一级频带分解的信号输出又作为下一级频带分解的输入,级数增加意味着级连的滤波器增多,造成的信号移位也增大。另一方面,由于每次小波系数分解/重构都要进行边界延拓,级数越多引起的边界失真也就越大。因此在实际应用中,确定分解/重构的级数要兼顾不同方面的影响。

8.6.3 小波零树编码[14]

零树编码是小波变换编码中很有代表性的算法。

如 3.5 节所述,在用二维 Mallat 算法对原始图像进行分解时,如分解为 M 层,则得到 3M 个高频子图(子带)和 1 个低频子图。用 L 表示一维低通滤波器,用 H 表示一维高通滤波器,则 LL,LH,HL 和 HH 表示 4 个具有不同频率特性和方向性的二维复合滤波器。LL 表示沿水

平方向和垂直方向作低通滤波，用于分解图像的低频分量，即原始图像的概貌；LH 表示沿水平方向作低通滤波，沿垂直方向作高通滤波，用于检测原始图像垂直方向的高频细节（如水平方向的边缘）。类似的，HL 用于检测原始图像水平方向的高频细节（如垂直方向的边缘），而 HH 则用于检测原始图像斜线方向的高频细节。图 8.10 表示对图像作 3 层小波分解的情况。低频子图保留了原始图像的主要特征，聚集了原始图像的绝大部分能量。各高频子图中大量的系数很小，对应原始图像的平缓变化区域；而少量的大幅值小波系数则对应着图像的细节，如边缘和纹理结构等，所以同一方向的各高频子图在视觉上具有相似的轮廓，即存在相似性。图像分解之后，低频子图与各高频子图的像素数量之和等于原始图像的像素数量，所以小波变换本身并不能压缩图像。

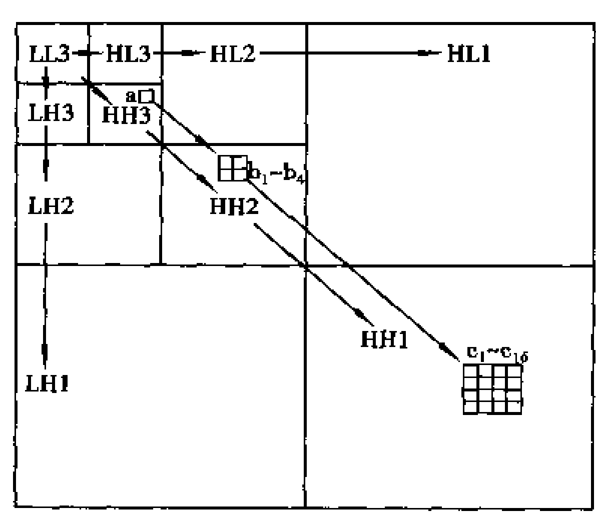


图 8.10　图像的小波分解

由于图像能量主要集中在低频子图系数和各高频子图的少量系数上，可以取一个适当的阈值 T，绝对值小于 T 的小波系数称为次要系数，否则称为重要系数。小波零树编码充分利用利用了各高频子图的相似性。较低分辨率上的系数称为父系数，与父系数同一方向但在较高分辨率上的系数称为其子系数。统计实验结果表明：如果某一父系数为次要系数，则与之对应的大片子系数均是次要系数的概率很大，这就是小波系数的零树特性。在图 8.10 中，如在 HH3 子图中位于 (i,j) 的系数 a 是次要系数，则可以认为在 HH2 子图中位于 $(2i,2j)$，$(2i+1,2j)$，$(2i,2j+1)$ 和 $(2i+1,2j+1)$ 的 4 个子系数 $b_1 \sim b_4$ 也是次要系数，进一步还可以把 HH1 子图中的 16 个子系数 $c_1 \sim c_{16}$ 也视为次要系数。在编码过程中，根据阈值 T 将系数 a 分为 4 类：

①a 及其各级子系数都是次要系数，称 a 为零树根。在后续的编码中，对零树根的子系数不予编码。而在解码时，将零树根及其各级子系数全置为零。

②a 是次要系数，但它的子系数中有 1 个或多个是重要系数，称 a 为孤立零。

③a 是正的重要系数。

④a 是负的重要系数。

在零树编码中，希望找到尽可能多的零树根以压缩编码，因此从最低分辨率开始按 Z 字形逐级向高分辨率扫描各个子图，如图 8.11 所示。此外，为了提高速度，对每个子图内的系数按 2×2，4×4，…的区域顺序扫描。判断系数类型的程序流程如图 8.12 所示，由于系数有 4 种类型，可用 2 比特标识符进行标识，在编码中将标识符保存下来，例如零树根、孤立零、正重要系数和负重要系数的标识符分别为 00、01、10 和 11。

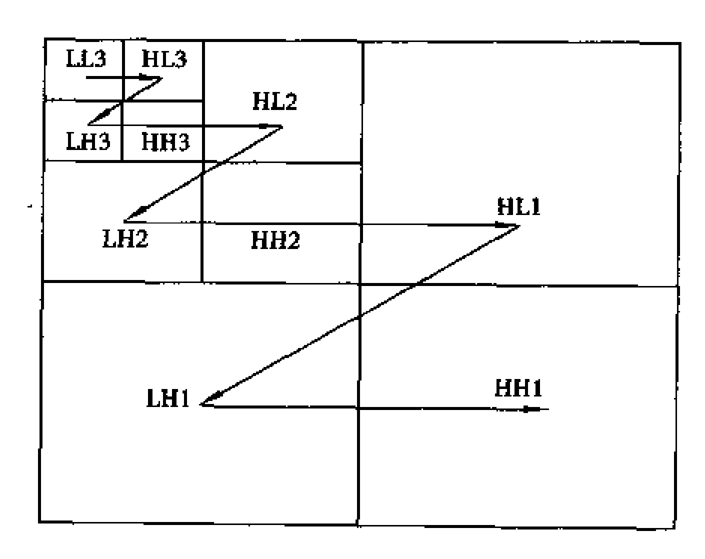


图 8.11 系数编码时的扫描顺序

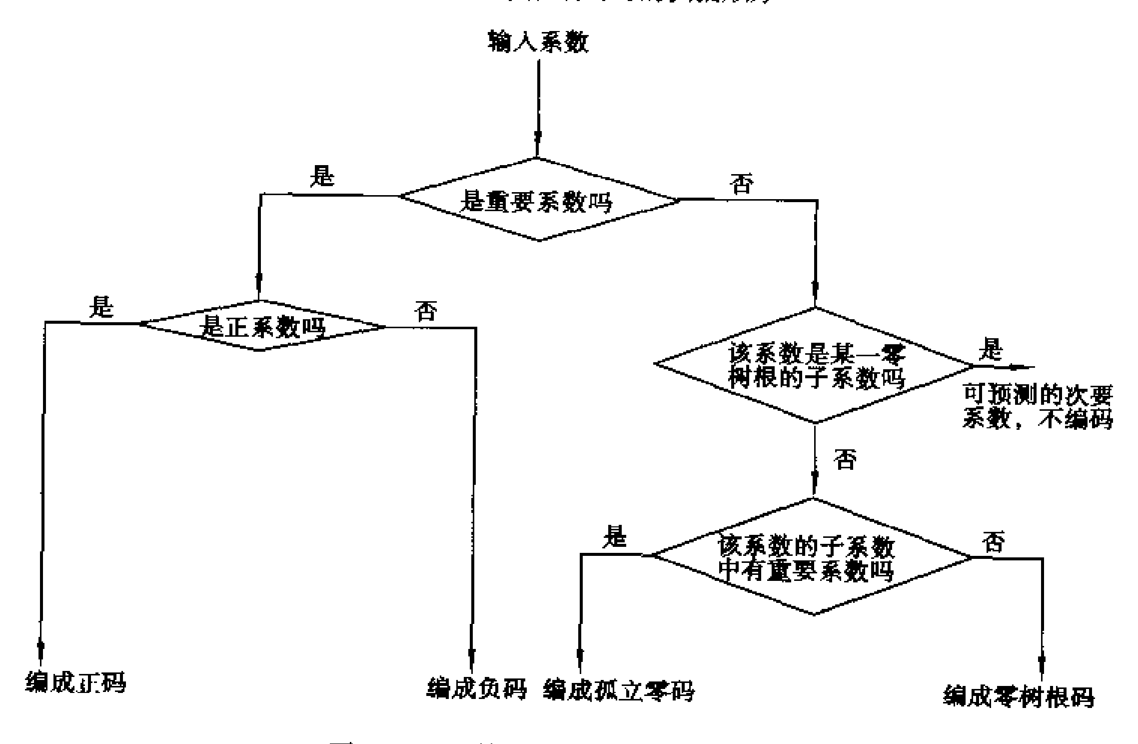


图 8.12 系数类型编码输出的程序流程

在零树编码中，还需要对正、负重要系数的幅值进行量化，如图 8.13 所示。为了减少存储的比特数，需要把重要系数的幅值编码与阈值联系起来，具体过程为：选择初始阈值 T 使之满足 $2T >$ 最大的小波系数，且 $T = 2^n$。用 T 对所有系数进行检测，检测到重要系数时，判断其所属区间（设区间长度为 L，则区间总数为 T/L），输出类型码和区间代码（其比特数为 $\log_2(T/L)$）；然后再用 $T/2$ 为阈值重复上述过程，以此类推，每次都使阈值减少一半，直到压缩比与图像质量达到要求为止。

在图像解码过程中，首先恢复最重要的系数，再恢复阈值减半后的系数，依次重复；也可以根据图像质量随时停止解码。解码时根据代码将该系数恢复成其所属区间的中间值。如采用的区间长度为 $L = 8$，则解码最大误差不超过 4。

在零树编码中，需要存储或传输的数据有类型码和区间代码两种数据流，可对它们进行熵编码以提高压缩比。

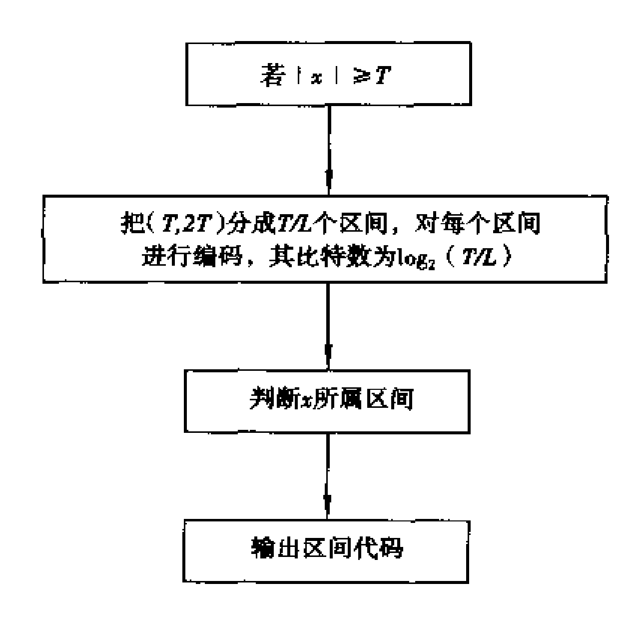


图 8.13　幅值编码流程图

第 9 章

泛函分析简介

9.1 集及其运算

所谓集合,是指具有某些共同性质的事物的全体。

例如:N(自然数集),Z(整数集),Q(有理数集),R(实数集),C(复数集),K(标量集)。标量集既可以是实数集,也可以是复数集。上面列举的是数集,实际上具有某些共同性质的所有函数可以构成函数集。

通常有两种描述集的方式:

1)列出集中所有的元素。例如:$N = \{1, 2, \cdots, n \cdots\}$

2)代表元素描述法。例如:$L^2(R) = \{x = f(t): \int_R |f(t)|^2 \mathrm{d}t < \infty\}$

上式中,$L^2(R)$是所有可测的、平方可积的函数集。x 是集的代表元素,它是一个函数(可以是一个实函数,也可以是一个复函数)。冒号后面的数学表达式进一步描述了代表元素的性质,即函数模的平方在整个实轴上的积分为一有限值。积分是勒贝格意义下的积分,而不是一般的黎曼积分。从信号处理的角度来看,函数 $f(t)$ 是一个一维信号,t 既可以是时间,也可以是空间;$f(t)$ 可能是随时间变化的电压或电流,也可能是随空间分布的光波的幅度。我们知道,功率与电压或电流的平方成正比,与光振动的幅度的平方成正比,故 $f(t)$ 平方的积分可理解为信号的能量。所以,所有能量有限的模拟信号都属于$L^2(R)$。

又例如:$l^2 = \{x = (\xi_l)_{l \in Z}: \ \xi_l \in K, \ \sum_l |\xi_l|^2 < \infty\}$

是所有平方可和的数列集。显然,所有能量有限的数字信号都属于 l^2。

定义了集之后,要讨论集和集之间的关系。

定义 9.1　映射:如图 9.1 所示,X 和 Y 为集,A 是 X 的子集,即 $A \subset X$。如对 A 中的任意元素 $x \in A$,均有 Y 中惟一的元素 y 与之对应,将这种对应关系称为从 A 到 Y 的映射 T,记为

$$y = Tx \tag{9.1.1}$$

y 称为 x 的像。A 称为 T 的定义域,记作 $D(T)$。所有像的集合称为 T 的值域或像集,记作

$R(T)$。映射 T 又称为从 X 到 Y 的算子。

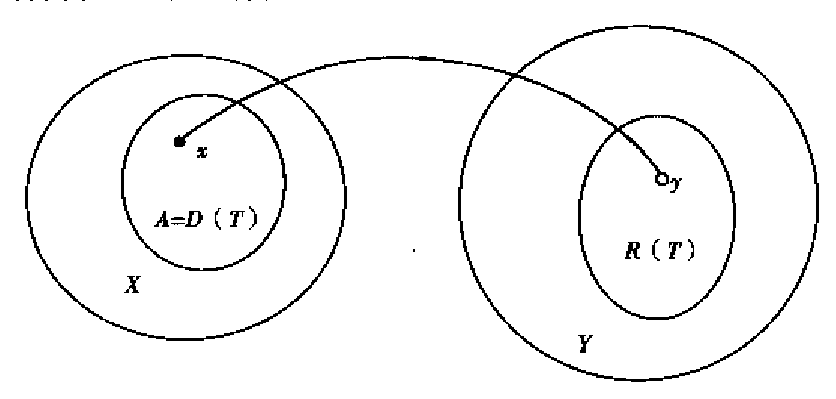


图9.1 映射

其实映射可以看成函数概念的推广。例如 $y=f(t)$ 是一个实函数，即 $t\in R$，且 $y\in R$，则函数实际上是从实数集 R 到实数集 R 的映射，函数是将一个实数映射为另一个实数。又如傅立叶变换，假定信号 $f(t)$ 的傅立叶变换为 $F(\omega)$，$f(t)\in L^2(R)$，根据巴塞瓦定理知 $F(\omega)\in L^2(R)$，则傅立叶变换是从 $L^2(R)$ 到 $L^2(R)$ 的映射。或者说，傅立叶变换是将一个函数映射为另一个函数。

映射实际上是某种数学运算。例如傅立叶变换，将函数 $f(t)$ 乘以 $e^{j\omega t}$ 的共轭再积分就得到它的傅立叶变换 $F(\omega)$。$L^2(R)$ 中任意两个元素的内积定义为

$$\langle f(t),g(t)\rangle = \int_R f(t)g^*(t)\mathrm{d}t$$

故傅立叶变换可表示为内积形式

$$F(\omega) = \int_R f(t)\mathrm{e}^{-j\omega t}\mathrm{d}t = \langle f(t),\mathrm{e}^{j\omega t}\rangle$$

9.2 距离空间

在集的基础上，引入一些拓扑结构可以形成不同的空间。距离是一种几何结构，它是通常欧几里德空间中距离概念的推广。

9.2.1 距离空间的定义

首先定义距离空间，然后列举几个重要的距离空间。

定义 9.2 设 X 是一个集，如能对 X 中任意两个元素 x 和 y 定义距离 $d(x,y)$，使其满足

1) $d(x,y)\geqslant 0$，d 是实的，有限的非负数 (9.2.1)

2) $d(x,y)=0$，当且仅当 $x=y$ (9.2.2)

3) $d(x,y)=d(y,x)$ （对称性） (9.2.3)

4) $d(x,y)\leqslant d(x,z)+d(z,y)$ （三角不等式） (9.2.4)

则称 X 是一个距离空间，记为 (X,d)，d 是定义于 X 上的距离。上列 4 式称为距离空间的公理。在集的基础上，先定义距离，然后证明它满足上列 4 条公理，则这个集才形成一个距离空间。

例 9.1 欧几里德平面 R^2（二维欧几里德空间）。

如图 9.2，欧几里德平面 R^2 的两个元素 x 和 y 是平面上的两个点。用两个有序的实数（横坐标和纵坐标）来描述平面上的一个点。定义

$$d(x,y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2}$$

则不难验证，它满足上述 4 个距离公理。故 (R^2,d) 为一距离空间。

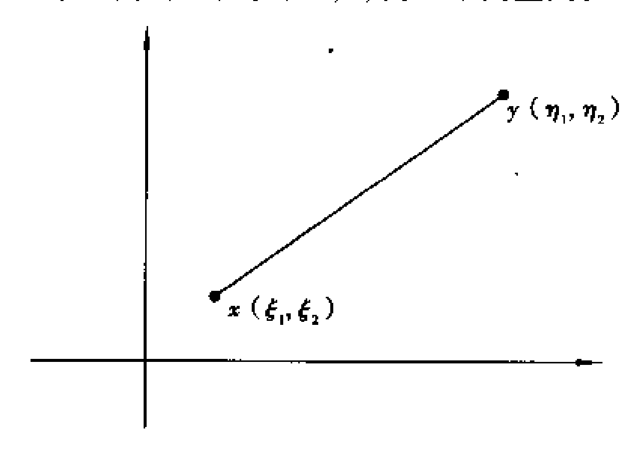


图 9.2 二维欧氏空间

还可以定义另一形式的距离

$$d'(x,y) = |\xi_1 - \eta_1| + |\xi_2 - \eta_2|$$

且也满足距离公理，故 (R^2,d') 也为一距离空间。从这个例子可见，对同一集，可定义不同的距离，从而构成不同的距离空间。

例 9.2 p 方可和序列集 l^p（$p \geqslant 1$，为一确定实数）。

p 方可和序列集 l^p 定义为

$$l^p = \{x = (\xi_l)_{l \in Z};\xi_l \in K, \sum_l |\xi_l|^p < \infty\}$$

如定义距离

$$d(x,y) = \left(\sum_{i=1}^{\infty} |\xi_i - \eta_i|^p\right)^{1/p}, \qquad x = \{\xi_i\}_{i=1}^{\infty}, y = \{\eta_i\}_{i=1}^{\infty}$$

则距离公理中的前 3 条很容易验证，第 4 条要利用 Hölder 不等式来验证。可见 p 方可和序列集 l^p 也可以构成一个距离空间。

例 9.3 平方可积函数集 $L^2(R)$。

定义 $L^2(R)$ 中任意两个元素 $f(t)$ 和 $g(t)$ 之间的距离为

$$\left\{\int_{-\infty}^{+\infty} |f(t) - g(t)|^2 \mathrm{d}t\right\}^{1/2}$$

可以验证它满足 4 条距离公理，所以平方可积函数集 $L^2(R)$ 构成一个距离空间。

9.2.2 极限，完备性

极限是数学分析中最基本的概念，微分和积分都是在极限概念的基础上定义的。如 $x \in R$，则 $\Delta x \to 0$ 实际上是指实轴上两个点 x 和 $x + \Delta x$ 之间的距离趋近于零。所以在距离空间中定义距离之后，也可以在距离空间中引入极限概念。

定义 9.3 距离空间的极限。

设$\{x_n\}$为距离空间(X,d)中的一个点列,若存在$x\in X$,使得

$$\lim_{n\to\infty} d(x_n,x) = 0 \tag{9.2.5}$$

则说$\{x_n\}$是收敛的,x称为$\{x_n\}$的极限,写成

$$\lim_{n\to\infty} x_n = x \quad 或 \quad x_n \to x \tag{9.2.6}$$

必须注意:不能把距离空间中的点简单地理解为欧氏空间的点。例如p方可和序列空间l^p中的一个点是一个序列。与l^p类似,也可以定义p方可积函数空间$L^p(R)$,它的点是一个函数。由于距离是一个非负实数,故(9.2.5)式中极限的概念可以按以前学过的极限概念理解,而(9.2.6)式应按(9.2.5)式理解,即x_n趋近于x应理解为$d(x_n,x)$趋近于零。

有了极限的概念之后,和数学分析中类似,可以引入映射连续性的定义。

定义 9.4 映射连续性。

设$X=(X,d)$和$Y=(Y,\tilde{d})$是两个距离空间,$T: X\to Y$是从X到Y的映射。如对X中的某一点$x_0\in X$及任意小的一个数$\varepsilon>0$,总存在$\delta>0$,对满足$d(x,x_0)<\delta$的所有$x\in X$,使$\tilde{d}(Tx,Tx_0)<\varepsilon$,则称映射$T: X\to Y$在点$x_0$连续。

如T在X内所有点连续,则称T为连续映射。

定义 9.5 柯西列。

设$\{x_n\}$是距离空间(X,d)中的一个点列,如对任意的$\varepsilon>0$,总存在$N(\varepsilon)$使得

$$d(x_m,x_n) < \varepsilon, \forall m,n > N \tag{9.2.7}$$

则$\{x_n\}$称为X的柯西列或基本列。

完备:若距离空间X中任意柯西列都收敛,则说X是完备的。

在实数域内,柯西列和收敛列是等价的。但对一般距离空间而言,收敛列比柯西列条件更强,即收敛列一定是柯西列,但柯西列却不一定是收敛列。

9.2.3 开集和闭集

定义 9.6 开球,邻域。

给定距离空间(X,d)中的一点$x_0\in X$及一实数$r>0$,称集

$$B(x_0,r) = \{x\in X: d(x_0,x_0) < r\}$$

为开球,x_0为球心,r为半径。

半径为ε,球心为x_0的开球$B(x_0,\varepsilon)$称为点x_0的邻域。

定义 9.7 内点,开集。

设M为距离空间(X,d)中的一个子集,若对M中的某一点$x_0\in M$,总存在$\delta>0$,使得$B(x_0,\delta)\subset M$,则称x_0为M的内点。

若M中所有的点均为其内点,则称M为开集。

读者如能回想一下实轴上开区间的概念,则对开集的理解会有一定帮助。

定义 9.8 极限点(聚点),闭集。

设M为距离空间(X,d)中的一个子集,x_0的任意邻域至少包含一个异于x_0且属于M的点x,即$x\neq x_0, x\in M$,则称x_0为M的极限点(或聚点)。

请注意:M的极限点可能属于M,也可能不属于M。

若M的所有极限点均属于M,则称M为闭集。

再次提醒注意：内点和极限点不是互相排斥的概念，实际上，内点必然是极限点。类似的，开集和闭集也不一定是两个互相排斥的概念。

9.2.4 闭包、稠密子集

它们是小波分析中常用到的两个重要概念。

定义 9.9 闭包。

由 M 及它所有的极限点组成的集称为 M 的闭包，表示为 $\overline{M}$。

闭包 $\overline{M}$ 是包含 M 的最小闭集。前面已经提请读者注意，M 的极限点可能属于 M，也可能不属于 M。显然，将不属于 M 的极限点扩充进来，就构成 M 的闭包。

定义 9.10 稠密子集。

设 X 为一距离空间，且 $A \subset B \subset X$。若对任意的 $x \in B$，均存在 $\{x_n\} \in A$ 当 $n \to \infty$ 时，$x_n \to x$，则称 A 在 B 中稠密，或称 A 是 B 的稠密子集。若 $B = X$，则称 A 在 B 中处处稠密。

可以证明：有理数集 Q 是实数集 R 的稠密子集，也就是说，任何一个无理数都可以用一个有理数列去无限逼近。无理数是一个无限不循环小数，可表示为 $x = a_0 . a_1 a_2 \cdots$，其中 a_0 为整数部分，a_1 为小数点后第 1 位，a_2 为小数点后第 2 位，a_n 为小数点后第 n 位。令有理数列 $\{x_n\}$ 的第 n 个数为 $a_n = a_0 . a_1 a_2 \cdots a_n$，当 $n \to \infty$ 时，显然有 $x_n \to x$。

闭包和稠密子集是有联系的，实际上有下述定理：距离空间 X 的子集 M 是稠密子集的充分必要条件是 $\overline{M} = X$。

可分距离空间：若距离空间 X 具有可数且稠密的子集，则称 X 是可分的。

9.3 线性空间

线性空间具有代数结构，主要是在集的基础上引入了加和数乘这两种代数运算。回想一下熟知的傅立叶变换，它将一个信号展开为若干正弦信号的线性叠加，其中就有数乘和加法运算。将小波函数伸缩和平移形成一个函数族，小波变换就是将信号展开为它的线性叠加。

9.3.1 线性空间（矢量空间）

线性空间通常又称为矢量空间，其元素称为矢量。

定义 9.11 线性空间。

对集 X，如可定义加和数乘运算，且满足如下公理

1）对 $\forall x, y \in X$ 及 $\forall a, b \in K$，使得

$$x + y \in X, ax \in X, 1x = x \quad \text{(可定义加和数乘运算)} \tag{9.3.1}$$

$$a(bx) = (ab)x, a(x + y) = ax + ay \tag{9.3.2}$$

2）总能定义零元素 $0 \in X$，使得 $0 + x = x$；总能定义负元素 $(-x) \in X$，使得 $x + (-x) = 0$。

3）对 $\forall x, y, z \in X$ 使得

$$x + y = y + x, \quad x + (y + z) = (x + y) + z \tag{9.3.3}$$

则称 X 为在数域 K 上的线性空间。

（9.3.1）式意味着可在线性空间中定义加和数乘运算，且任意两个矢量 x 和 y 之和 $x + y$

仍然属于该线性空间;线性空间中任意一个矢量 x 乘以数 a 得到的矢量 ax 仍然属于该线性空间。这两点非常重要。(9.3.2)式意味着线性空间必须满足乘法结合律和乘法对加法的分配律。另外,(9.3.3)式告诉我们,线性空间中的加法必须服从加法交换律和结合律。

9.3.2　线性闭包

定义 9.12　线性子空间。

设 Y 是线性空间 X 的非空子集,如对任意的 $y_1,y_2\in Y$ 以及任意的标量 a_1,a_2 均有 $a_1y_1+a_2y_2\in Y$,则称 Y 为 X 的线性子空间。

定义 9.13　线性闭包。

设 M 是线性空间 X 的任意非空子集,M 中矢量所有可能的线性组合之集 Y 称为 M 张成的线性子空间。如将 M 中矢量线性组合的极限也包含到 Y 中,则 Y 称为 M 张成的线性闭包。

在小波多分辨率分析中,尺度函数 $\phi(t)$ 经二进伸缩和平移得到的函数族 $\{\phi_{j,n}(t)\}_{n\in Z}$ 是 $V_j\subset L^2(R)$ 的正交规范基,V_j 中的任意函数可表示为 $\{\phi_{j,n}(t)\}_{n\in Z}$ 的线性组合或线性组合的极限,故 V_j 是 $\{\phi_{j,n}(t)\}_{n\in Z}$ 张成的线性闭包。表示为 $V_j=\overline{\mathrm{span}\{\phi_{j,n}(t)\}_{n\in Z}}$。

又例如:三维欧氏空间 R^3 中的两个矢量(1,0,0)和(0,1,0)是 R^3 的一个非空子集,由它张成的线性子空间就是平面 R^2。

9.4　赋范空间,巴拿赫空间

范数是欧氏空间中矢量长度概念的推广。在线性空间的基础上引入范数这样一个几何结构,就构成赋范空间。赋范空间一定可以构成距离空间,可见它是一种性质很好的空间。也就是说,赋范空间既是线性空间(可定义数乘和加法),又是距离空间(可定义距离)。

9.4.1　赋范空间

定义 9.14　赋范空间。

设 X 是一个线性空间,如对任意的 $x\in X$ 可定义一个范数 $\|x\|$,且满足

1) $\|x\|\geqslant 0$,即范数是一个非负实数　　(9.4.1)

2) $\|x\|=0$,当且仅当 $x=0$　　(9.4.2)

3) $\forall a\in K,\ \|ax\|=|a|\cdot\|x\|$　　(9.4.3)

4) $\|x+y\|\leqslant\|x\|+\|y\|$　　(三角不等式)　　(9.4.4)

则称 X 为赋范空间。

对赋范空间 X,可根据范数定义如下形式的距离

$$d(x,y)\triangleq\|x-y\|,\ \forall x,y\in X \tag{9.4.5}$$

很显然,$d(x,y)$ 满足距离公理(9.2.1)式~(9.2.4)式,故 X 构成距离空间 (X,d),而(9.4.5)式称为由范数导出的距离。

完备的赋范空间称为 Banach 空间。下面举几个巴拿赫空间的例子。

例 9.4　n 维欧几里德空间 R^n 及酉空间 C^n,定义

$$\|x\| = \left(\sum_{i=1}^{n} |\xi_i|^2\right)^{1/2} \tag{9.4.6}$$

例9.5 p 方可和数列空间 l^p，定义

$$\|x\| = \left(\sum_{i} |\xi_i|^p\right)^{1/p} \tag{9.4.7}$$

例9.6 p 方可积函数空间 $L^p[a,b]$

$$L^p[a,b] = \left\{x = f(t) : \int_a^b |f(t)|^p \mathrm{d}t < \infty\right\} \tag{9.4.8}$$

定义范数为

$$\|x\| = \left(\int_a^b |f(t)|^p \mathrm{d}t\right)^{1/p} \tag{9.4.9}$$

对于平方可积函数空间 $L^2(R)$，其任意矢量 $f(t)$ 的范数定义为

$$\|x\| = \left(\int_{-\infty}^{+\infty} |f(t)|^2 \mathrm{d}t\right)^{1/2} \tag{9.4.10}$$

也就是说，范数的平方就是信号的能量。在信号分析中，往往用一个函数 $g(t)$ 去逼近信号 $f(t)$，则 $g(t)$ 逼近 $f(t)$ 的均方根误差其实就是 $f(t)$ 和 $g(t)$ 的距离，或者说 $f(t)-g(t)$ 的范数，即

$$\|f(t) - g(t)\| = \left(\int_{-\infty}^{+\infty} |f(t) - g(t)|^2 \mathrm{d}t\right)^{1/2} \tag{9.4.11}$$

例9.7 有界数列空间 l^∞

$$l^\infty = \{x = \{\xi_i\}_{i=1}^{\infty} : |\xi_i| < C_x, C_x \in R\} \tag{9.4.12}$$

$$\|x\| = \sup_i |\xi_i| \tag{9.4.13}$$

例9.8 连续函数空间 $C[a,b]$。

它是区间 $[a,b]$ 上所有连续函数的集，其上的范数可定义为

$$\|x\| = \max_{t\in[a,b]} |f(t)| \tag{9.4.14}$$

可以验证，以上各例中所定义的范数均满足范数公理，所以它们都是赋范空间。

9.4.2 线性算子

定义9.15 线性算子。

线性算子是满足如下两个条件的映射：

1）映射 T 的定义域 $D(T)$ 是一个线性空间，值域 $R(T)$ 是在同一数域上的线性空间。

2）$\forall x,y \in D(T)$，$a \in K$，

$$T(x+y) = Tx + Ty, \quad T(ax) = aTx \tag{9.4.15}$$

等式左边的加和数乘是在 $D(T)$ 上定义的，而右边的加和数乘是在 $R(T)$ 上定义的。

如映射 T 连续，则称其为连续算子。

回忆一下映射的定义，若 X 和 Y 是数量空间，那么 T 就是一般意义下的函数。若 X 为赋范空间而 Y 为数量空间，那么 T 称为泛函。实际上，范数本身就是一个泛函。

9.4.3 有界线性算子

定义9.16 有界线性算子。

令 X 和 Y 均为赋范空间，$T: D(T)\to Y$，若 $\forall x \in D(T)$ 均有一实数 C 使得

$$\frac{\|Tx\|}{\|x\|} \leq C, \quad \forall x \in D(T) \subset X \tag{9.4.16}$$

则称 T 为有界线性算子。

注意：$\|x\|$是在 X 上定义的范数，而$\|Tx\|$是在 Y 上定义的范数。

容易证明：连续算子是有界算子，反过来也成立。假若 T 为有界算子，则当$\|x-x_0\|\to 0$时，由(9.4.16)式知$\|Tx-Tx_0\|\to 0$，所以 T 必然是连续算子。

能否找到一个最小的实数使(9.4.16)式成立呢？由(9.4.16)式可得

$$\frac{\|Tx\|}{\|x\|} \leq C, \quad \forall x \in D(T) - \{0\} \tag{9.4.17}$$

可见最小的 C 就是$\frac{\|Tx\|}{\|x\|}$，$\forall x \in D(T) - \{0\}$的最大值，称为有界线性算子的范数(或模)。

$$\|T\| = \sup_{x \in D(T)} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x \in D(T) \\ \|x\|=1}} \|Tx\| \tag{9.4.18}$$

显然，

$$\|Tx\| \leq \|T\| \cdot \|x\| \tag{9.4.19}$$

上式在泛函分析中经常使用，请读者留意。

算子范数的几何意义：通过映射 T，矢量 x 的"长度"变换为它的像长度$\|Tx\|$，故$\frac{\|Tx\|}{\|x\|}$表示通过映射之后，某个矢量 x 的伸缩系数。对不同的矢量，伸缩系数可能不同，而算子范数$\|T\|$就是最大的伸缩系数。

9.5　内积空间，希尔伯特空间

内积是欧氏空间中两个矢量点积概念的推广。内积空间中的投影定理在信号处理中应用十分广泛。内积空间可以构成赋范空间。

9.5.1　内积空间

定义 9.17　内积空间。

设 X 是一线性空间，若对其中任意两个矢量 x 和 y 可定义内积$\langle x,y\rangle$，且满足

1)$\langle x,y\rangle$为一实数或复数，若内积为实数，$x\neq 0$时，$\langle x,x\rangle$大于0　(9.5.1)

2)$\langle x,y\rangle = \langle y,x\rangle^*$　(9.5.2)

3)$\langle ax,y\rangle = a\langle x,y\rangle$，$a\in K$　(9.5.3)

4)$\langle x,y+z\rangle = \langle x,y\rangle + \langle x,z\rangle$　(9.5.4)

5)$\langle x,x\rangle = 0$，当且仅当 $x=0$　(9.5.5)

则称 X 为内积空间。

对内积空间 X，可按内积定义范数

$$\|x\| = \sqrt{\langle x,x\rangle} \tag{9.5.6}$$

且满足(9.4.1)式～(9.4.4)式，故 X 可构成赋范空间。由(9.5.6)式定义的范数称为由内积

导出的范数。

由内积空间可构成赋范空间,但并非所有的赋范空间均可构成内积空间。实际上,l^p 和 $L^p[a,b]$ 是赋范空间,当且仅当 $p=2$ 时,它们才可构成内积空间。

下面举几个内积空间的例子。

例 9.9 n 维欧几里德空间 R^n 及酉空间 C^n,定义内积

$$\langle x,y\rangle = \xi_1\eta_1^* + \xi_2\eta_2^* + \cdots + \xi_n\eta_n^* \tag{9.5.7}$$

且很容易证明它满足内积公理(9.5.1)式~(9.5.5)式,所以 R^n 和 C^n 是内积空间。对三维欧几里德空间 R^3,内积实际上就是两个矢量的点积。

例 9.10 平方可和数列空间 l^2,定义内积

$$\langle x,y\rangle = \sum_{i=1}^{\infty} \xi_i\eta_i^* \tag{9.5.8}$$

且很容易证明它满足内积公理,所以 l^2 是内积空间。

例 9.11 平方可积函数空间 $L^2[a,b]$,定义内积

$$\langle x,y\rangle = \int_a^b x(t)y^*(t)\,\mathrm{d}t \tag{9.5.9}$$

且很容易证明它满足内积公理,所以 $L^2[a,b]$ 是内积空间。当区间 $[a,b]$ 扩展为整个实轴时,可知 $L^2(R)$ 是内积空间。

完备的内积空间称为希尔伯特空间。可以证明:$R^n,C^n,l^2,L^2[a,b]$ 均为希尔伯特空间。

9.5.2 同构算子与伴随算子

定理 9.1 Riesz 表示定理:如果 T 是 Hilbert 空间 X 上的有界线性泛函,则存在惟一的 $x_T\in X$,对任意的 $x\in X$,有

$$Tx = \langle x,x_T\rangle \tag{9.5.10}$$

定义 9.18 如果 $\forall x_1,x_2\in X$,有 $\langle Tx_1,Tx_2\rangle=\langle x_1,x_2\rangle$,则称 T 为同构算子。

定义 9.19 算子 T 为 X 到 Y 上的有界算子,算子 T 的伴随算子 T^* 定义为 Y 到 X 上的算子,且满足

$$\langle x,T^*y\rangle = \langle Tx,y\rangle,\quad \forall x\in X, y\in Y \tag{9.5.11}$$

对于伴随算子 T^*,其模(范数)满足

$$\|T^*\| = \|T\|, \|T^*T\| = \|T\|^2 \tag{9.5.12}$$

定义 9.20 如 T 是从 Hilbert 空间 X 到 X 上的算子,且 $T=T^*$,则称 T 是自伴算子。

9.5.3 正交性

定义 9.21 两个矢量的正交。

对 $x,y\in X$,若 $\langle x,y\rangle=0$,称 x 和 y 正交,记为 $x\perp y$。

根据两个矢量正交的定义,可以进一步定义矢量与子集及子集与子集的正交:M 和 N 为内积空间 X 的两个子集。若某一 $x\in X$ 与 M 中的任意矢量正交,则称 x 与 M 正交,记为 $x\perp M$。若 N 中所有的矢量均与 M 正交,则称 N 与 M 正交,记为 $N\perp M$。

于是可以定义正交 ,它在小波分析中有非常重要的应用。

定义 9.22 正交补。

集$\{x:x\perp M,x\in X\}$称为M在X中的正交补，记为$M^{\perp}$。

也就是说，M的正交补$M^{\perp}$就是X中所有与M正交的矢量集，故$M^{\perp}\perp M$。

关于正交性，还有如下重要定理：

1）勾股定理：若$x\perp y$，则$\|x+y\|^2=\|x\|^2+\|y\|^2$。

2）设L为内积空间X的稠密子集，若$x\perp L$，则$x=0$。

也就是说，除零矢量外，没有任何一个矢量和稠密子集正交。

3）内积空间X的子集M的正交补$M^{\perp}$是X中的一个闭线性子空间。

9.5.4　投影定理

定义 9.23　直和。

设X为线性空间，Y和Z是其两个线性子空间。若对X中的任意矢量$x\in X$，均有惟一的表达式$x=y+z,y\in Y,z\in Z$，则说X是Y和Z的直和，记为$X=Y\oplus Z$。

可以说，泛函分析在信号处理中最重要的应用之一就是投影定理。

定理 9.2　投影定理：

设M是内积空间X的完备线性子空间，则X可表示为M及其正交补的直和，即

$$X=M\oplus M^{\perp}\tag{9.5.13}$$

也就是说，X中的任意矢量x可惟一的表示为

$$x=x_0+x_{\perp},x_0\in M,x_{\perp}\in M^{\perp}\tag{9.5.14}$$

x_0称为x在M上的正交投影，上式称为x的正交分解。

投影定理可以形象地表示于图9.3，以便于理解。

由（9.5.14）式可以看到，$x-x_0=x_{\perp}$，$x-x_0\perp M$。很容易证明

$$\|x-x_0\|\leqslant\|x-x_m\|,\forall x_m\in M\tag{9.5.15}$$

由于$(x_0-x_m)\in M$，所以$(x-x_0)\perp(x_0-x_m)$，根据勾股定理，有

$$\|x-x_m\|^2=\|(x-x_0)+(x_0-x_m)\|^2=\|x-x_0\|^2+\|x_0-x_m\|^2\geqslant\|(x-x_0)\|^2$$

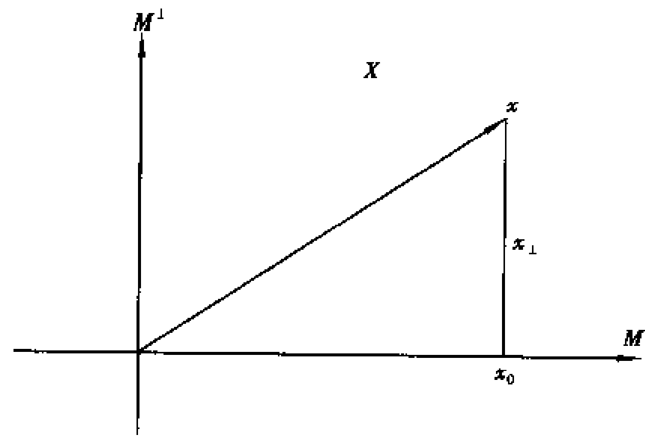


图9.3　投影定理

这就证明了（9.5.15）式。该式说明，x到其正交投影的距离是x与M中所有点之间的距离中最小的，$\|x-x_0\|$称为x到M的距离。换句话说，如用M中的矢量去逼近x，则正交投影x_0是最佳逼近元，这就是所谓的变分原理。

实际上，（9.5.14）式定义了如下映射

$$P:X\to M,x\mapsto x_0=p_x\tag{9.5.16}$$

P称为x到M的投影算子。可以证明，P是一个有界算子。

9.5.5　正交分解

对于有限维空间，例如我们最熟悉的三维欧几里德空间R^3。沿x,y,z 3个坐标轴的单位矢量$\vec{i}=(1,0,0)$，$\vec{j}=(0,1,0)$，$\vec{k}=(0,0,1)$

两两正交，且它们各自的范数都等于 1，那么 R^3 中的任意矢量 $\vec{r}$ 可分解为

$$\vec{r} = x\vec{i} + y\vec{j} + z\vec{k} \tag{9.5.17}$$

其中

$$x = \vec{r}\cdot\vec{i}, y = \vec{r}\cdot\vec{j}, z = \vec{r}\cdot\vec{k} \tag{9.5.18}$$

分别是 $\vec{r}$ 的 3 个坐标值。称 $\{\vec{i},\vec{j},\vec{k}\}$ 是 R^3 的正交规范基。

定义 9.24 正交规范基。

设 $\{e_n\}$ 是内积空间 U 中的矢量族，若满足

$$\langle e_i, e_j\rangle = \begin{cases}0, & i \neq j\\ 1, & i = j\end{cases} \tag{9.5.19}$$

则称 $\{e_n\}$ 是 U 的正交规范基。

定义 9.25 广义傅立叶级数。

设 $\{e_n\}$ 是内积空间 U 的正交规范基，对 $\forall x \in U$，称

$$c_n = \langle x, e_n\rangle \tag{9.5.20}$$

为 x 关于 $\{e_n\}$ 的傅立叶系数，并称

$$x = \sum_n c_n e_n = \sum_n \langle x, e_n\rangle e_n \tag{9.5.21}$$

为 x 关于 $\{e_n\}$ 的傅立叶级数。

例 9.12 傅立叶级数。

不难验证，$\{e^{in\omega_0 t}/\sqrt{T_0}\}_{n\in Z}, \omega_0 = 2\pi/T_0$ 是 $L^2[0, T_0]$ 的正交规范基，即

$$\langle e^{in\omega_0 t}/\sqrt{T_0}, e^{im\omega_0 t}/\sqrt{T_0}\rangle = \frac{1}{T_0}\int_0^{T_0} e^{i(n-m)\omega_0 t}dt = \begin{cases}0, & n \neq m\\ 1, & n = m\end{cases}$$

这意味着任意属于 $L^2[0, T_0]$ 的函数 $f(t)$ 均可展开为

$$f(t) = \frac{1}{\sqrt{T_0}}\sum_{n=-\infty}^{+\infty} c_n e^{in\omega_0 t}$$

它就是我们熟悉的傅立叶级数，其中

$$c_n = \langle f(t), e^{in\omega_0 t}/\sqrt{T_0}\rangle = \frac{1}{\sqrt{T_0}}\int_0^{T_0} f(t)e^{-in\omega_0 t}dt$$

投影定理告诉我们，对任意 $x \in X$，它在 $M \subset X$ 上的正交投影 x_0 是对 x 的最佳逼近，下述定理告诉我们如何求取 x_0。

定理 9.3 设 $M = \mathrm{span}\{e_n\}$ 为由正交规范矢量族 $\{e_n\}$ 张成的 X 的线性子空间，$\forall x \in X$，

$$x_0 = \sum_n \langle x, e_n\rangle e_n \tag{9.5.22}$$

是 x 在 M 上的正交投影，且 x_0 是对 x 的最佳逼近。证毕。

证明 设 x 属于 X 但不属于 M，由(9.5.22)式有

$$x - x_0 = x - \sum_n \langle x, e_n\rangle e_n$$

对 $\{e_n\}$ 中的任意矢量 e_m，

$$\begin{aligned}\langle x - x_0, e_m\rangle &= \langle x, e_m\rangle - \sum_n \langle x, e_n\rangle\langle e_n, e_m\rangle\\ &= \langle x, e_m\rangle - \langle x, e_m\rangle = 0\end{aligned}$$

由于 M 是 $\{e_n\}$ 张成的线性子空间,故上式意味着 $x-x_0$ 与 M 中的所有矢量正交,从而 $x-x_0$ 与 M 正交,这就证明了 x_0 是 x 在 M 上的正交投影。证毕。

上述定理证明,一旦有了 X 的线性子空间 M 的正交规范基,就很容易求取任意 $x\in X$ 在 M 上的正交投影。在小波多分辨率分析中,尺度函数二进伸缩和平移形成的正交规范基 $\{\phi_{j,n}\}_{n\in Z}$张成了 $L^2(R)$的一系列子空间 V_j,信号 $f(t)\in L^2(R)$在 V_j 上的正交投影可以写成正交规范基的线性组合。

例9.13 取 $L^2[0,T_0]$的正交规范基的前 N 项 $\{e^{in\omega_0 t}/\sqrt{T_0}\}_{-N\leqslant n\leqslant N}$,它张成 $L^2[0,T_0]$的一个线性子空间,任意属于 $L^2[0,T_0]$的函数 $f(t)$在其上的正交投影

$$f_0(t)=\frac{1}{\sqrt{T_0}}\sum_{n=-N}^{+N}c_n e^{in\omega_0 t},c_n=\langle f(t),e^{in\omega_0 t}/\sqrt{T_0}\rangle$$

是对 $f(t)$的最佳逼近,逼近的均方误差为

$$\|f(t)-f_0(t)\|^2=\frac{1}{T_0}\sum_{\substack{n=-\infty\\ n\notin[-N,N]}}^{+\infty}|c_n|^2$$

在傅立叶分析的工程应用中,可以取前 N 次谐波,N 越大,逼近误差越小。

对其他内积空间,也可以用类似方法进行分析。

参考文献

[1] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. IEEE Trans. IT-36(5):961-1005,1990

[2] I. Daubechies. Ten lectures on wavelets. SIAM, Philadelphia, 1992

[3] S. Mallat. A Wavelet Tour of Signal Processing, Second Edition. Academic Press. 1998

[4] S. Mallat. A theory of multiresolution signal decomposition: The wavelet transform. IEEE Trans. PAMI-11(7):674-693, 1989

[5] S. Mallat. Multifrequency Channel Decomposition of Images and Wavelet Models. IEEE Trans. ASSP-37(12):2091-2110, 1989

[6] S. Mallat and W. L. Hwang. Singularity Detection and Processing with Wavelets. IEEE Trans. IT-38(2):617-643, 1992

[7] S. Mallat and S. Zhong. Characterization of Signals from Multiscale Edges. IEEE Trans. PAMI-14(7):710-732, 1992

[8] I. Daubechies. Orthonormal bases of compactly supported wavelets. Commun. on Pure and Appl. Math., 41:909-996, November 1988

[9] 李水根,吴纪桃.分形与小波.北京:科学出版社, 2002

[10] 陈武凡等.小波分析及其在图像处理中的应用. 北京:科学出版社,2002

[11] 杨福生.小波变换的工程分析与应用. 北京:科学出版社,1999

[12] 赵松年,熊小芸.子波变换与子波分析. 电子工业出版社,1996

[13] 程正兴.小波分析算法与应用. 西安交通大学出版社,1998

[14] 李弼程等.小波分析及其应用. 电子工业出版社,2003

[15] M. Antonini, et al. Image Coding Using Wavelet Transform. IEEE Trans. Image Proc.-1(2):205-220, April 1992

[16] Wim Sweldens. The Lifting Scheme: A Construction of Second Generation Wavelets. SIAM J. Math. Vol.29, No.2, pp.511-546, March 1998